DOCTOR OF PHILOSOPHY

Returns to Education
Evidence from the UK

Xu, Lei

Award date:
2018

Returns to Education: Evidence from the UK

Lei Xu

Supervised by Prof Yu Zhu and Dr Paul Seaman

Submitted to

School of Business

at

University of Dundee

In partial fulfilment of the requirements for the degree of

Doctor of Philosophy

2018

Acknowledgements

I am honoured to learn and work with my supervisors, Prof Yu Zhu and Dr Paul Seaman. I am mostly indebted to Prof Yu Zhu for the support and all the comments and feedbacks to my thesis. I also have benefited from Paul's incisive criticism and the comments to my English.

I would like to thank all the seminar participants at the University of Dundee the 2016 and 2017 participants at the Work Pensions and Labour Economics (WPEG) annual conference, the 2016 Scottish Economic Society conference, 2017 Scottish Graduate Programme in Economics (SGEP) Residential conference, 2017 European Association of Labour Economists (EALE).

I am very grateful for the SGPE scholarship for these three years, the Royal Economic Society (RES) conference grant, and the University of Dundee conference grant.

Finally, I am also greatly indebted to my parents for their continuous support to my study.

Declaration

I hereby declare that the work in this thesis was carried out in accordance with the Regulations of the University of Dundee. I have not used any published materials in this thesis. Three empirical chapters are single authored.

"Wage Penalty of Vocational Education: Evidence from the UK (Chapter 3)", presented to the 2016 Scottish Economic Society conference.

"Does age-dependent minimum wage affect employment? evidence from UK (Chapter 4)", presented to the 2016 Work Pensions and Labour Economics (WPEG) annual conference, the 2017 Scottish Graduate Programme in Economics (SGEP) Residential conference, 2017 European Association of Labour Economists (EALE).

"Quantitative effects of higher education expansion on the returns to education: Evidence from the UK (Chapter 5)", presented to the 2017 Work Pensions and Labour Economics (WPEG) annual conference.

Abstract

This thesis consists of three independent chapters which address separate questions in relation to labor economics and economics of education

Wage Penalty of Vocational Education: Evidence from the UK (Chapter 3)

In this chapter, I examine the difference in wages between academic and vocational education in the UK based on Quarterly Labor Force Survey (QLFS) from 2001 to 2013. First I examine the crude wage differences between vocational and academic education. To further test the differences between two types of education, I examine the effect of the education expansion on the returns to higher levels of vocational qualification, based on the difference-in-difference (DID) methodology. The results suggest that the reform has negative effects on the wage of holders of higher levels of vocational education. The penalties vary considerably, depending on the type of vocational qualification.

Does age-dependent minimum wage affect employment? evidence from UK (Chapter 4)

The chapter studies the age-dependent minimum wage in the UK, which is used to regulate the flow of young workers into the labor market. In this chapter, I examine the employment effect of becoming eligible for higher minimum wage rate by applying Regression Discontinuity (RD). The results suggest that an increase in the minimum wage has a positive effect on employment probability for higher skilled worker covered by the minimum wage but not for lower skilled workers, and it may also lead to crowding out effect coming from higher skilled workers. Moreover, higher skilled workers tend to transfer from a temporary job into a formal job more easily after becoming eligible for higher minimum wage rate and this pattern is the opposite for lower skilled workers. The evidence suggests that the labor market in which the minimum wage prevails is very competitive during the recession and lower skilled workers may bear the cost of competition due to the discontinuity caused by age-related increases in minimum wage.

Quantitative effects of higher education expansion on the returns to education: Evidence from the UK (Chapter 5)

This chapter studies the effect of the education expansion on the returns to education based on Quarterly Labour Force Survey (QLFS) and Understanding Society. After examining the heterogeneous returns, I apply the difference-in-difference (DID) methodology to examine the effect of the education reform on the returns and the matching Difference-in-Difference (MDID) methodology to account for the compositional change across cohorts since those newly recruited university graduates after the reform might be different from the previous graduate cohorts. Newly recruited university graduates consist of "fresh students" who entered universities as school leavers typically with A-level and workers with several years of work experience called "mature students". The MDID results show that the expansion of higher education mostly reduces the returns for fresh students and it mostly appears in the post-expansion period. The mature students have more stable returns compared with fresh students.

Contents

Chapter 1
General introduction

The economics of education is to study how the education has proceeded and the impacts of the education on social and economic outcomes. Education provides pivotal skills and helps individual get into employment. It is strongly correlated with employment, earnings, life satisfaction, and so on.

The discussion of the human capital is the core of the economics of education, especially after Gary Becker who is normally considered as the founder of the modern economics of education. After both extensive and intensive development of the human capital theory, we know that the human capital can increase a person's productivity directly and signal a person's potential productivity. Moreover, it is multi-dimensional and can also increase the adaptability in a changing environment. The human capital is distinctive by the types of education, innate abilities, quality of education, and other types of training. The pre-labour market factors have an influence on the individual's educational attainment, such as parental background, financial budget constraint, neighbourhood, and so on. Moreover, the regulations and the ideology of the education system are different across the world. It also deeply affects the individual's choices of education. The interventions are designed to achieve the potential productivity of a person and to meet the employer's needs of the skills.

In this thesis, I address three separate questions in relation to the outcomes of education.

The return to vocational education (Chapter 3)

In the UK, students will choose to continue their education or finding a job after the compulsory education at 16. Compared to the compulsory schooling, there are many types of qualifications from 16-19, broadly categorized as vocational or academic. Vocational education was rapidly developing in the UK. It is considered as an efficient pathway for people to acquire more skills. Finding the subtle differences between academic and vocational education is important regarding the political interest.

In this chapter, I examine the wage differences between academic and vocational education in the UK based on Quarterly Labour Force Survey (QLFS) from 2001 to 2013. First, I estimate the crude wage differences between vocational and academic education over time. To further test the differences between two types of education, I estimate the effect of the "Education Reform Act 1988" on the return to higher levels of vocational qualification, based on the difference-in-difference (DID) methodology. The results suggest that the reform has negative effects on the higher levels of vocational education. The negative effects vary considerably, depending on the type of vocational qualification.

Does age-dependent minimum wage affect employment? Evidence from the UK (Chapter 4)

The average years of education increased significantly in these three decades in the UK. There are still a considerable amount of individuals who only have GCSEs in the labour market. They may be young or have relative few working experiences, making them the most vulnerable employees in the labour market. This chapter studies the age-dependent minimum wage in the UK, which is used to regulate the flow of young workers into the labour market. I examine the employment effect of becoming eligible for higher minimum wage rate by applying Regression Discontinuity (RD). There is a strong evidence of heterogeneous effects by qualifications. The results suggest that there is no significant effect of the increased minimum wage on the employment probability for individuals below GCSE and the significant discontinuity mainly comes from higher skilled workers. Individuals with higher numbers or grades of GCSE have a higher probability of being employed after increasing the minimum wage. There are no significant effects for individuals with lower numbers or grades of GCSE. The number of GCSEs held by employed individuals and proportion of 5+ GCSEs among employed workers increase across the threshold, suggesting that there is a crowding out effect. Moreover, the results strongly suggest that higher-skilled workers tend to find a more full-time permanent or full-time job after becoming eligible for higher minimum wage rate. But on the other hand, lower skilled workers have a lower probability of finding a full-time permanent job. The evidence of higher employment probability and higher satisfying job accession probability may imply that there is a crowding out effect coming from higher skilled workers. Although the result is modestly significant, the higher number of GCSEs and the higher proportion of 5+ GCSE suggest the existence of crowding out effect directly.

Quantitative effects of the higher education expansion on the returns: Evidence from the UK (Chapter 5)

The university graduates increased significantly after the Education Reform Act 1988. It provided more opportunities for students pursuing a degree. It also opened doors for people who have working experiences or fewer years of fundamental education compared to the students who finish A-level. In the UK, there are a considerable amount of students who start a job prior to obtaining the highest qualification, known as “returning students”. Few pieces of evidence have been found regarding the return to returning students. In this chapter, I apply the relatively new methodologies to re-examine the effect of the education reform on the returns to academic education and highlight the returns of those returning students.

After examining the heterogeneous returns, I apply the difference-in-difference (DID) methodology to examine the effect of the education reform on returns and the matching Difference-in-Difference (MDID) methodology to account for the ability bias since those newly recruited university graduates after the reform might be different from the previous graduates. Newly recruited graduates consist of “fresh students” who graduate from A-level or other schools and workers with several years of working experiences known as the “returning students”. I notice that there is a compositional change among graduates. It is expected that the degree of relaxation of requirements changed over time. The proportion of the returning students among those university graduates increased significantly after 1976. The compositional change may bring uncertain heterogeneities into the results.

From the DID results, both fresh students and returning students don’t have a significantly negative effect in the pre-expansion period in which there are increasing the supply of university graduates who don’t have working experiences before. The MDID results correct the innate ability bias and estimate the quantitative effect on returns for returning students, leading to a negative effect to the returning students. Interestingly, during the post-expansion period, the DID results show that fresh students have been more affected by the increasing supply of returning students. After correcting the ability bias, both results become more negative.

The thesis is organized as follows. The second chapter introduces the education system in the England and methodologies in this thesis. The third chapter is the first empirical chapter which examines the differences between academic and vocational education. The fourth chapter is the second chapter which examines the effects of increasing minimum wage on employment. The fifth chapter examines the effect of education expansion on the returns to education. The last chapter is the conclusion which describes the main findings of this thesis and the future research plan.

Chapter 2

The UK education system and review of the methodologies

2.1 Review of Further Education in England.

In England, children enter primary school at the beginning of the school year when they turn into five years old. The schools normally start from September. That means the primary schools will receive the students following their fourth birthday.[1] The leaving age of compulsory education is now 18 since summer 2015, arising from 17 in 2013 and 16 before 2013. The education choices are flexible between 16 and 18. Students may select from full-time education, work-based learning, or part-time education and part-time employed. In the following, I will focus on the education system which is effective till 2013 and covers the whole sample periods from 2008 to 2012 inclusive.

The outputs at each Key Stage are set out in the National Curriculum and the National Tests would be taken at the end of each stage, age 7 (Key Stage 1), age 11 (Key Stage 2), age 14 (Key Stage 3), age 16 (Key Stage 4). Students are expected to meet certain levels accordingly at the end of each stage's learning, primary school (Entry Level), lower secondary school (Level 2), upper secondary

[1] See https://www.citizensadvice.org.uk/family/education/school-education/access-to-education/

school (Level 3). The secondary education usually takes place between 11-18. Until 2013, at the end of compulsory schooling (secondary schooling), all students are expected to take at least five General Certificate of Secondary Education (GCSE) (Hupkau et al, 2016).[2] In 1988, the O-level and CSE exams are combined into General Certificates of Secondary Education (GCSEs).[3] It gives students wider options regarding the subjects and also encourages more students who swing between O-level and other qualification into education. The students will receive scores from A-G, where A is the highest score in GCSEs.[4] After the transition from GCSEs to O-level, the A-C scores in the GCSEs will give students more influence in the labour market. GCSE belongs to the National Qualification Framework (NQF) and a GCSE with grades A-C belongs to Level 2 and a GCSE with grades D-G belongs to Level 1. The vocational GCSEs are introduced since 2000.

After Secondary schooling, some students who are normally between 16 and 19 continue to take further education. In England, Further Education (FE) refers to

---

[2] See Hupkau et al (2016) which describe the UK education system in a figure for the most intuitive understanding that how different levels of education are organized.

[3] The Act enacted in 1988, introduced major changes into the education system in England. First, the National Curriculum and the standard attainment test (SATS) for children 7, 11, 14 were introduced. Second, in order to provide better quality of education, the government adopted the concept of "quasi-market". The schools were encouraged to compete with each other and recruit students without fewer limits in order to increase the efficiency in the market. The schools had to find ways for the "consumers", namely the parents of students, to easily access the information to make a correct move. The results of this policy make some schools larger and others face the possibility of shutting down.

[4] A* was granted in 1994 to the top student in GCSEs.

the education after the compulsory education of 16.[5] It is a crucial part of the education system and has been developing over-time. There is no unified classification of Further Education. Broadly it could be grouped into three categories (Cuddy and Leney, 2005). The first category includes general tertiary colleges. Those institutions provide professional training in a wide range of subjects. The second category comprises sixth-form colleges that typically provide education for the 16 to 18 years old. Those schools specialize in providing full-time courses for both academic and vocational education. The third category includes private trainer providers. There are vast numbers of private education providers. Each specializes in a field. Compared to academic education, vocational education serves a versatile role in the labour market. Vocational education can provide more specific skills to help the unemployed workers after a job-loss or to enhance their skills in the existing workplace or obtain more education in order to pursue a higher education. Accordingly, there are wide ranges of institutions to serve people with different purposes. There are 1168 publicly funded Further Education institutions in the UK in 2014, including 247 tertiary colleges, 94 Sixth Form colleges, and 827 private training providers respectively (Hupkau and Ventura, 2017). They also conclude that the population of students at Further Education institutions has decreased by one million over the two years after 2005 due to the decrease in the value of the certificate in the

[5] Here I only summarize the education system in England since it contains most students and Education systems in Wales and North Ireland are rather similar to England.

labour market. The funding that supported the Further Education institutions previously is transferred to support the learners to pursue Full Level qualifications.[6][7]

2.1.1 Academic track.

After compulsory education, students may continue to stay in an academic track education, a vocational track education or a mixture of the two. Large numbers of the student who remain in the academic-track education will choose Advanced-Levels between 16 to 18 years old. They will be rewarded the GCE Advanced Subsidiary (AS) qualification after 1 year and the A2 Level after two years at Key Stage 5, Level 3. Normally they will take three or more subjects over 2 years and those are the standard requirements for university entrance. Most of the A-Level graduates will continue their education in universities. The new A-Level system allows them to take half of the exams in the first year. In the second year, they will be awarded the full GCE A-Level after completing the

[6] The General National Vocational Qualifications (GNVQs) were introduced in 1992. It is designed to be more classroom-based taught vocational qualifications. Compared to National Vocational Qualifications (NVQs) that are more work-related, GNVQs are more flexible. Due to the structure of GNVQs, students may continue higher levels vocational education or go to workplace instead. The National Qualifications Framework (NQF) was introduced in 2000 in England, Wales and Northern Ireland. All Vocational Education and Training (VET) were regulated under this framework and made comparable to other qualifications. The NQF specifies nine levels (entry level and levels 1 to 8). Each level consists of different qualifications, including both academic and vocational qualifications. It is ended in 2007.

[7] The Department of Education has recently created a set of non-A-Level categories at Level 3, known as “Applied Generals and Tech Levels” and a set of non-GCSE categories at Level 2, known as “Tech Certificates” (Hupkau et al, 2016).

exams successfully. The subjects in A-Level have great influence on choosing the subjects in the university. For example, math and science subjects are essential in many universities for a science related degree.

The major changes for higher education system started from 1988 to 1994, followed by the “Further and Higher Education Act 1992”, which converted some higher education institutions and polytechnics into universities. This changed pushed more students to become university graduates directly. Moreover, it also relaxed the requirements for students because of the large reduction of the funding from the government. Both changes resulted in the large increase in the university graduates.

2.1.2. Vocational track.

For those students who are unwilling to stay on academic education, some of them will stay on the vocational education. The vocational education system in the UK is developing over time. Back in the 1960s, after the Industrial Training Act 1964, the vocational education normally consisted of one day a week of study at a further education college, along with an apprenticeship, mostly provided by City and Guild of London Institute (CGLI). This system didn’t work well in practice and it quickly collapsed during the 1970s to 1980s (Machin and Vignoles, 2005). Similar to CGLI, Business Education Council (BEC) and Technician

Education Council (TEC) also provided the day-release courses in the 1970s.

The Business and Technology Education Council (BTECs) was found in 1983 after merged by the BEC and TEC.[8] The BTECs are regulated by the Qualifications and Curriculum Authority (QCA) in the same way as NVQs. The aim of the BTECs is to provide courses in business and technology for the 16-19 years old. It was designed to benefit the student by providing both general and vocational education. It offers courses to meet the need of industries in a range of subjects, including Ordinary Certificate/ Diploma (ONCs/ONDs) and Higher Certificate/ Diploma (HNCs/ONDs).

The National Council for Vocational Qualifications (NCVQ) was set up by the government in 1986 (now known as the Qualifications and Curriculum Authority) and the role of NCVQ is to accredit the National Vocational Qualifications (NVQs). The NVQs aim to increase the skills for those in work. There are leading bodies that are responsible for organizing the connection between employers and education providers. The NVQs consists of eight levels. The highest level could compare to the postgraduate degree. There are many units of competency tests for each NVQ and the tests will cover the knowledge applied to specific job functions. The assessments of the NVQs normally carry two aspects. One includes the performance of the practical works. The other includes the traditional skill tests and questioning. There are three main NVQs providers: the further

[8] See West and Steedman (2003) for more details.

education sector colleges, private training organizations and some companies. One of the keys for providing NVQ is to set up a network with companies. This leads to one of the major merits of the NVQs that it provides work-related content in a simulated work environment and it aims to provide both the core-skills and the work-related skills. Some barriers have been proposed when organizing the NVQs, such as small companies unwilling to provide positions, the cost of organizing assessment and training supervisors. More importantly, there is a lack of the connection of the needs of the skills between students and employers' need. And the assessment is more arbitrary compared to normal education.[9]

2.1.3 Modern Apprenticeship.

The modern apprenticeship was introduced in 1993 by the Conservative Government. It was designed to fill the skill gap between the employees and the employers by providing more flexible and work-related knowledge, mainly for 16-24 years-old workers with intermediate skills. In principle, it is open to workers regardless their age. The apprentices who are above 25 years old are regarded as the older-apprenticeship. The fields of the modern apprenticeship not only cover the traditional craft and manufacturing, but also the service sectors. Another major change was the financial support of the government and

9 See the Further Education Funding Council (1994) for more details.

the way of organization of the funding to the apprenticeship.[10] The cost of the training now is transferred to the government via the Skills Funding Agency.[11] Compared to the previous youth training schemes which deliver the skills mainly at Level 1 and 2, the modern apprenticeship covers wider workers including Level 3 and Level 4 (Fuller, 2016). The modern apprenticeship in England, supported by the government starts from Level 2 (broadly equivalent to GCSEs at grades A* to C), known as Intermediate Apprenticeship. The higher apprenticeship includes Level 3, known as Advanced Apprenticeship and Level 4, known as Higher Apprenticeship.[12] Compared to the old apprenticeship, the new modern apprenticeship includes more industries and frames its structure into NQF (Fuller and Unwin, 2003). Due to the nature of the structure of apprenticeship, it is quite flexible across different apprenticeships, in terms of pay, length, off-the-job training, etc. In order to strengthen the training and creditability of the apprenticeship, Qualifications and Curriculum Authority (QCA) developed the Technical Certificates which includes the details of knowledge training, assessment and teaching program (Fuller and Unwin, 2003). Concerns have been raised recently in relation to the effectiveness of apprenticeship, the probability of accrediting the skills rather than providing the skills to workers.

[10] In 2014, each apprentice received 1500 pounds supported by the government and the organizations didn't need to pay the national insurance fees for the apprentices in 2016 (Fuller, 2016).

[11] The Skills Funding Agency will be responsible for registering the training providers rather than the employers.

[12] The Level 4 apprenticeship was introduced in 2009-2010.

2.1.4 Education finance.

After completing compulsory education at 16, many students will continue their study either in an academic track or in a vocational track. Students from lower social-economic groups are less likely to participate in the post-16 education. One of the barriers is the financial constraint. To encourage the less advantaged students to continue their study, the Department for Education and Employment (DfEE) introduced the Education Maintenance Allowances (EMA) in 2004.[13] It covers all students in the United Kingdom. The 16-19 Bursary Fund was introduced to replace the EMA in England in 2011. Both the EMA and 16-19 Bursary Fund are designed to provide financial support[14] for students who stay in full-time education between 16-18 years old. The magnitude of the support is dependent on the annual income of the family.

In England, there had been no tuition fees for university students since 1962. However, as a part of the Teaching and Higher Education Act 1998, tuition fees were reintroduced at 1000 pounds per year in 1998. Together with the introduction of the tuition fees, the maintenance grants were also abolished (Dearden et al, 2011).[15] At the meantime, the government replaced the grants with tuition loans and maintenance loans in order to offset the negative impact

[13] Before the formal EMA was introduced in 2004, there is a pilot provision from 1999 to 2003 in 15 Local Education Authorities (Dearden et al, 2001).
[14] The funds are expected to spend on clothing, food, transport, book, and the equipment in the class, see https://www.gov.uk/1619-bursary-fund
[15] They also pointed out that the real value of grants had been largely eroding in the 1990s.

on students. Many factors are contributive to the introduction of the tuition fees. One of the most important reasons could be the increasing numbers of university students and the decreasing average funds per person, especially due to the education expansion after 1988. Moreover, the flexible tuition fees may also increase the competition among universities on the basis of the marketization-oriented concept.

Later on, the tuition fees have been raised up to 3000 pounds in 2004 for the new students after 2006-2007, arising out of the Higher Education Act 2004 (Dearden et al, 2012). During those periods, the proportion of university students from the low social-economic groups was stagnant compared to the fact that the total numbers of university graduates increased significantly. In order to encourage more students with disadvantaged family background into universities, the maintenance grants were reintroduced at the same time at 1040 pounds per year and the students can obtain up to 2700 pounds at most (Dearden et al, 2011). Another major change as a result of the Act was that all fee loans could be repaid after graduation. All of the fees, grants, and loans are dependent on the basis of the parental income. In 2012, it has been largely raised again up to 9000 pounds per year.[16]

2.2 Methodologies Reviews.

[16] See House of Commons (2016).

One of the main features that differ micro-econometrics from statistics lies in the distinction in the ideology between "relation" and "causal relation". The empirical studies of labour economics focus on estimating the "treatment effect" or the "causal relation". People are interested in the counterfactual outcomes before they take the move.[17] It involves the "what if" question, leading to the popularity of the concept, known as "potential outcome" which are the imaginary outcomes when the individual takes both choices. It requires an unrealistic parallel world. That's why the applied economists focus on eliminating the selection bias, caused by the differences in characteristics between treatment takers and non-takers. It can easily be tackled by the social experiment. However, since social experiments are rarely existing, empirical studies rely on constructing the "counterfactual group" in most of the time, called the "natural experiment method".[18]

2.2.1 Instrument Strategy and Control Function method.

The regression is the cornerstone of applied econometrics nowadays as well as the increasing popularity of the reduced form. The Ordinary Least Squared (OLS) is a standard method to examine a simple relationship between the variable of interest and the outcome on the basis of observable covariates. The basic idea is

[17] The so-called "frequentist inference" assumes that the estimates remain unchanged across individuals. The Bayesian inference assumes that the estimates follow the random process under a probability distribution.

[18] Blundell et al (2009) use "natural experiment method" to denote a methodology that constructs a comparison group in a properly designed experiment.

to adjust the observable covariates between the treatment and the control group to make them identical, like a random social experiment. It is a method that belongs to "selection on observables".

Under the framework of potential outcomes, the causal effect is identified on the basis of two parallel groups. A regression on observables can rarely meet the Conditional Independence Assumption (CIA), in which it assumes that the variable of interest is independent of the potential outcome. In another word, the variable of interest is randomly assigned given the control variables. If the CIA fails, the independents are correlated with the residuals, known as "endogeneity". There are three main sources of endogeneity, Omitted Variable Bias (OVB), Reverse Causality and Measurement Error. The bias caused by omitting variables for which need to control is called the "Omitted Variable Bias (OVB)". For most empirical labour studies, the variables are insufficient to capture all potential channels (lower R-squared) and some channels are closely related to the variable of interest such that it is needed to control for (collinearity). Due to the data generating process of the repeated cross-section data, the problem of endogeneity and limited information can't be avoided. Moreover, we don't even know the extent to which the OVB and endogeneity exist in the research.

In order to bypass the endogenous problem in the regression, Instrument Variable (IV) strategy is widely used by the applied economists. The choice of instruments is the pivotal issue in an IV strategy. Contrary to adding more variables into the regression, the IV approach tackles this problem in an opposite direction in which it claims the exogeneity of the instruments and the variation of the variable of interest brought by the instruments is used to identify the causal effect between the outcome and the variable of interest. The regression of the endogenous variable on the instruments is the "First Stage" and then the regression of the instruments and other exogenous variables on the outcome is called the "Reduced From".[19] The estimate of an IV strategy is the ratio of the reduced form to the first stage. A good instrument needs to meet two restrictions. First, the instrument needs to have a strong and clear relationship with the endogenous variable or the variable of interest. Second, the instruments can only affect the outcome in the way through the first stage. It suggests that the instruments are only correlated with the variable of interest rather than other determinants of the dependent variable or the error terms, known as the "exclusion restriction".[20] The exclusion restriction doesn't guarantee the functioning of IV strategy. The instruments may have insufficient variation to capture the change in the outcome in the first stage, leading to the "weak instruments". This will happen if there is no selection on the idiosyncratic gain. In addition, a good instrument that meets the exclusion restriction rule is hard to

[19] The exogenous covariates should be included in both first stage and reduced form. The failure of inclusion of the covariates would incur the biased result due to the fact that the covariates are not independent with the residuals in the first stage. If the covariates are sufficiently independent from the instruments, including the covariates would produce more precise results. See Angrist and Pischke (2009).

[20] The treatment effect might be heterogeneous. A heterogeneous effect may help us clarify the distinction between internal validity and external validity (Angrist and Pischke, 2009).

find in practice.

Very similar to IV strategy, the endogenous variable in the reduced form can be replaced by the estimated value in the first stage. The parameter of the estimated variable is called the estimator of the Two-Stage-Least-Squared (TSLS). Compared to the normal IV strategy, TSLS can include more instruments in the first stage when the instruments capture the same causal effect. When implementing the TSLS, it is essential to keep the same covariates in both first stage and reduced form (Angrist and Pischke, 2009).

The causal effect of a treatment is identified when comparing the same treatment group and control group. This derives the terms such as "Always-takers", "Compliers", "Never-takers", and "Defiers" under the context of Angrist and Pischke (2009).[21] In practical studies, a researcher needs to be careful to clarify that the causal effect on which subgroup is identified.[22] Different instruments estimate different causal parameters and a set of instruments may trigger more people caused by a single instrument. On the basis of the conceptual framework of potential outcomes, researchers are interested in estimating the average treatment effect (ATE). The ATE only exists in a perfectly randomized experiment, in which the treatment is randomly assigned across

[21] Always-takers are defined as those who receive the instrument and take the treatment. Never-takers refer to those people who receive the instrument and don't take the treatment. Compliers refer to people who take the treatment whenever they receive the instrument. Defiers refer to people who don't take the treatment due to the instrument.

[22] The estimated effect would be different with including different samples.

observations. It equals to average treatment effect on the treated (ATT) plus the selection bias caused by the differences between the treatment group and the control group. In the IV strategy, the ATT is estimated on the basis of the random assignment as an instrument for the treatment received. Due to the fact that the treatments normally follow a self-selection process even when the treatments are randomly distributed since individuals will choose if they will pick up the lotteries.[23] If the compliance rate is not perfect, then it drives another concept, known as the "Intention-to-treat" effect (ITT). It occurs when individuals select the treatment given their own needs. That also differentiates between those who have the lotteries and turn it down and those who have the lotteries and pick it up. In other words, we should be careful when we draw the conclusions which target to which groups of people.

Moreover, similar to TSLS, a control function method has been proposed based on the instruments. The bias comes from the situation where the independents are correlated with the residuals. Compared to TSLS, the control function method includes the residuals that are estimated by regressing the endogenous variable on the instruments. The endogenous variable would become independent after including residuals, served as a control variable in the reduced form. It can also allow for the heterogeneous effect by including interaction terms between the residuals and endogenous variable, called "Correlated Random Coefficient (CRC)".

[23] $E[Y_i|D_i = 1] - E[Y_i|D_i = 1] = E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1] + E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]$

In practice, the control function method produces similar results to the TSLS.[24]

In general, OVB can be regarded as selection bias from the perspective that the sample selection bias could be solved by adding more variables into the regression. Selection bias mostly refers to self-selection of individuals to the samples in practical scenarios. The difference is that OVB can be transferred into sample selection bias. Perhaps the most famous method to tackle the selection bias is the Heckman Model (Heckman, 1979). It is widely known that the selection bias would be a big threat to the results. There is considerable literature focusing on "selection on un-observables" in the recent decades. I explain in more detail in the sensitivity test part.

One of the major limitations of the normal IV strategy is the assumption that the treatment effects are identical across treated observations. In a more general case, one should notice that the treatment effects are heterogeneous. Imbens and Angrist (1994) proposed the Local Average Treatment Effect (LATE) to identify the impact of treatment around the neighbour's characteristics when the impact of the instrument is highly heterogeneous. The LATE captures the difference in means for those individuals who switch to take the treatment, triggered by the instruments.[25] The LATE relies on four assumptions, independence, exclusion restriction, monotonicity and first stage. With these four assumptions, LATE will capture the effect of the treatment on the subpopulation that changes the

[24] See Wooldridge (2014), page 421.

[25] $LATE = \frac{E(Y|Z=z_1) - E(y|Z=z_2)}{Pr(D=1|Z=z_1) - Pr(D=1|Z=z_2)}$

treatment status due to the seemingly unrelated instrument, known as "Compliers". In practice, LATE is usually performed on the basis of TSLS. It estimates the weighted average of the causal effect on subgroups that are triggered by the instruments in the context of heterogeneity. TSLS estimates the weighted average of the causal effect of a subpopulation, identified by the instruments.[26]

Rather than identifying the compliers, Heckman and Vytlacil (2005) propose the Marginal Treatment Effect (MTE) to estimate the heterogeneous treatment effect.[27] Compared to the LATE that estimates the average treatment effect around a point in the distribution of unobserved factors (compliers), the MTE estimates the treatment effect on the basis of individuals' "net willingness-to-pay" or the distribution of unobserved factors.[28]

2.2.2 Matching.

Matching belongs to the non-parametric estimation. The goal of matching is to

---

26 See Angrist and Pischke (2009).

27 $\Delta^{MTE}(x, u_d) = E(Y_1 - Y_0 \mid X = x, U_d = u_d)$, where u_d is the mean net utility given covariates that can be estimated as propensity scores. This framework suggests that the MTE is a willingness-to-pay measure that the estimate is the marginal gain under the covariates on the basis of the u_d. If the MTE does not depend on u_d, then IV equals to MTE.

28 Triggered by the instrument the unobserved factors of the compliers lie in around a point in the distribution of unobserved factors. The unobserved distribution is estimated given the framework of propensity scores.

construct a control group that is to a large extent close to the treatment group on the basis of a set of observables. It makes covariate specific comparison between treated and non-treated observations and sums the weighted cells together to get the average treatment effect. It is a method that belongs to the data preprocessing rather than a statistical method.

Several assumptions are needed when applying the matching strategy. One identification assumption is known as "Unconfoundedness", selection on observables, or CIA. The "Unconfoundedness" implies that the potential outcome is independent of treatment variable given a set of observable covariates. Another is called "Overlap" or common support. It guarantees that observations with certain covariates can be found in both treatment and control groups. In reality, the common support could be a problem if the proportion of uncovered individuals is high to some extent. The results would be biased if the dropped observations are not randomly assigned too. The sensitivity tests would be needed in order to check the robustness of the results. With the spirit of matching strategy, several matching algorithms have been proposed, such as the Nearest Neighbor Matching, Caliper and Radius Matching, Kernel Matching, etc.

There are also some restrictions for the matching strategy. First of all, the major weakness is the selection on observables. That means that researchers are required to have all the relevant information in order to satisfy the CIA. Second, matching is rather data-hunger and the numbers of observations are not enough in a survey dataset sometimes. Third, the results exist arbitrary when choosing the balance between the quantity of information and the proportion of common support (Blundell and Costa Dias, 2009).

As an empirical strategy, the results would be volatile given the different process of implementing the strategy. Before performing the matching strategy, one should access the overlap and trim the original data on the basis of t statistics or normalized differences. The traditional rule is to drop those treated observations for which there are no on-support observations in the control group (Robin and Rosenbaum, 1983). Huber et al (2013) propose a procedure of dropping observations in which those non-treated observations with a higher importance and treated observations with propensity scores above a threshold. Given their procedures, Lechner and Strittmatter (2017) suggest that the threshold could be at the maximum or the 99% quantile of the propensity score in the non-treated subpopulation. Second, one should also check the "selection on observables", although it is not testable directly. Imbens (2015) argues that the calculation in order to check the "Unconfoundedness" focus on estimating the effect of the treatment on a pseudo outcome, a variable which is unaffected by the treatment or the effect of the treatment on the lagged outcome. Lastly, after accessing the plausibility of the assumption of the matching strategy, one can estimate the causal effect of the treatment on outcomes.[29]

[29] Imbens (2015) argues that one can use replacement and bias-adjustment to reduce the bias.

2.2.2.1. Propensity Score Matching (PSM).

One of the limitations of a matching strategy is the large dimensionality, causing "curse of dimensionality". The curse of dimensionality is a phrase used in many subfields related to statistics. Here it refers to a situation that it is unlikely to find matched observations between treated group and control group in a matching strategy. Rosenbaum and Robin (1983) proposed to use a balancing score to perform a matching strategy. The scores have been estimated based on given observables, called "Propensity Scores". It becomes very popular because of its simplicity. The Propensity Scores are pivotal to the results. Compared to the full covariates matching strategy, the PSM only includes the information before the treatment in order to remove the unbalance. This will lead to higher asymptotic standard errors.[30] It converts the attention of the researchers from estimating the outcomes to treatment assignment, playing a flexible role together with other estimating strategies. Most empirical papers use a parametric model together with a matching. Although it is simple, the results could be rather volatile given the estimated propensity scores. Several alternatives have been proposed by applying non-parametric models or the relatively new machine learning technics. The next step is to perform matching by using the propensity scores rather than covariates after deciding the neighbours for each matched observation in the

[30] Angrist and Hahn (2004) argue that even though the asymptotic standard error can not be improved by PSM, there could be a gain in the precision in finite samples.

treatment group. In order to adjust the differences given the propensity scores, one can weight the treatment group and the control group to construct more balanced groups, or one can divide the sample into subsamples, called "blocking", or one can do the regression on the propensity scores directly. The way of matching and the selection of the bandwidth is normally an empirical question.

2.2.2.2. Coarsened Exact Matching (CEM).

The weights are the pivotal issue in a weighting strategy. Since the reweighting and matching type of strategy is more or less like a black box, we should be very careful when choosing the weights. The covariate matching and PSM still have the weakness. Making some covariates balanced is easy, but sometimes it also makes others even more unbalanced (Iacus et al, 2012). It is hard to choose the balance between covariates and matching quality, known as "equal percent bias reducing" (EPBR). CEM is a recently proposed method to help derive the weights and check the results from the PSM.

Basically, the CEM[31] is to coarsen each variable in order to group those substantively indistinguishable observations and then weights are estimated on the basis of the coarsened data given existing algorithms as well as the pruning of unmatched observations. The set of strata generated by the first step of the CEM is used to eliminate substantive differences across variables on the basis of

[31] The CEM belongs to the MIB (Monotonic Imbalance Bounding).

the chosen level defined by the coarsening. From my own perspective, the CEM is ideal for matching the variables which include dummies or discrete variables in an ambiguous way, such as including measurement error.

According to the explanation of Iacus's paper, CEM requires no assumption in relation to the data generating process and the imbalance after matching will not be larger than without matching, which is a unique feature compared to other matching strategies.

2.2.2.3. Matching vs. Regression.

Both regression and matching are based on constructing the counter-factual untreated group to estimate the treatment effect. They are the methods of selection on observables, all relying on the CIA. So there are no major differences between matching and regression from an empirical perspective.[32] The difference between matching and regression is the weights to sum the covariate-specific effects into a single value. A regression uses a variance weighted average treatment effect while a matching uses the probability of taking a treatment as weights. In other words, a regression should be same as a matching if the treatment is independent of covariates (Angrist and Pischke,

[32] Compared to regression, matching may lose information after those unmatched cases are discarded even when there are "good" reasons to drop those unmatched observations.

2009).[33]

Both regression and matching involve a certain amount of extrapolation. When implementing a matching strategy, a researcher needs to select bandwidth to determine the common support.[34] A regression that is not saturated and fails to meet the requirement of common support may affect the results by extrapolation. Compared to a matching strategy, there is also no absolute rule of dropping "no common support" or "thin common support"[35] in a regression, because the parametric regression will extrapolate the counterfactual outcome for the treated observations on the basis of the covariates. Although there is a fact that the empirical research is heavily based on the regressions, the common support has been neglected to some extent. But without the common support, the extrapolation over un-observables requires the absence of endogenous variables. On the other hand, matching doesn't need the exogeneity assumption compared to parametric regressions, given that the variables used for matching are determined before the treatment (Blundell and Costa Dias, 2009).

2.2.2.4. Weighting with Regression.

[33] Matching puts the most weights on the cells which there are most treated observations. Regression puts the most weights on the cells which the treated observations are the same as the control observations.

[34] Picking the bandwidth would be arbitrary since if the bandwidth is too small, then there are few observations in each cell, leading to dropping of the cells and the loss of information. On the other hand, if the bandwidth is too large, then there are more risks that treated group and control group are not identical.

[35] "No common support" refers to the situation that there is no observation in the control group that is similar to the observations in the treatment group.

In the standard econometric textbook, weighting can be used to tackle with the heterogeneity problem in a regression. But it is not very popular in micro-econometrics since the problem would not be serious when heterogeneity is mild (Angrist and Pischke, 2009). There are several uses for matching in a regression. In principle, weighting can be used to make your sample more close to the target population. One important use of weighting is to balance the sample when the sample is unrepresentative, for which it can be achieved by weighing with inverse probability weighting. Another interest regarding weighting is its practical use in estimating the causal effect. One popular way to construct a comparable treatment group and a control group is to implement the inverse probability weighting on the basis of propensity scores.[36] Rightful using of the reweighting can reduce the bias in a regression, but it may also increase the random error and bias the estimates. Implementing reweighting could be both empirical and hard.

2.2.3 Difference-in-Difference (DID) and Matching Difference-in-Difference (MDID).

Since it is costly to conduct a social experiment, researchers mostly take

[36] $ATE = E[Y(1) - Y(0)] = E\left[\frac{Y_i * T}{e(X_i)} - \frac{Y_i - (1-T)}{1-e(X_i)}\right]$. See Hirano and Imbens (2001); Angrist and Rokkanen, 2015.

advantage of the existing social and economic reforms to estimate the causal relations. After the work by Ashenfelter and Card (1985), the DID has been added into the toolkit. A mixed effect between the treatment effect and the common trend that is the default effect if there is no treatment is estimated by the subtraction between the outcomes of treatment group across the time t. The common trend can be estimated by the subtraction of control group. Then the causal effect can be estimated by the subtraction of the differences between treatment group and control group.

In a more general setting, the group and the period can be extended into multiple. Panel data is often used to apply the DID. One of the advantages of the panel data is that one can apply Fixed Effect (FE) model together with DID in order to remove the unobserved group or time effects. The main advantage of DID is that after the first difference in periods, it rules out the selection on an untreated outcome. The time period normally is one dimension, but it may refer to another dimension rather than the "time".

But the DID also relies on a strict assumption that the common trend would be identical to the treatment group. In practice, the common trend assumption can be easily violated either due to the fact that a control group will also be affected by the treatment or the control group is rather different from treatment group, leading to a failure of the common support.[37] The results would vary

[37] Abadie (2005) proposed a two-step procedure to deal with the problem that the treatment group and control group don't follow parallel paths. The STATA code calls

considerably with the control group selected by the researcher. Second, the results would be biased if the treatment has been anticipated. Third, the results would be biased due to the serial correlation problem (Bertrand et al, 2004). Much of the debate focuses on the validity of the randomness of the treatment assignment and its estimation format. Fourth, the results would be biased if the un-observables of individual are correlated with the choice of taking treatment. It guarantees that there is no compositional change across groups after the treatment has been implemented. By combining the merits of both matching and DID, the methods MDID may correct the bias from DID.[38] In other words, it is ideal to tackle the compositional change across either treatment group or control group due to the treatment.

2.2.4 Regression Discontinuity (RD).

The RD is a quasi-experimental design in which the probability of receiving treatment will be changed discontinuously across the threshold as a function of control variables. There are two types of RD, sharp and fuzzy. The sharp RD is when the treatment status has discontinuously changed after the threshold. The fuzzy RD is when the probability of taking treatment has discontinuously changed after the threshold. The essence of RD is to compare the treatment group on the right-hand side of the cut-off point and the comparison group on

"absdid", written by Kenneth Houngbedji.

38 See Heckman et al (1997).

the left side of the cut-off point. The control group from marginally below the threshold is valid counterfactual for treatment group from marginally above threshold (Hahn et al, 2001). Around the threshold, the treatment status is independent of all variables no matter observable and unobservable like a random assignment. So the causal effect is identified by the jump around the threshold. The discontinuity of a RD strategy can be regarded as a weighted average treatment effect on all individuals in the presence of heterogeneous treatment effects (Lee and Lemieux, 2009). Another major merit of RD is the graphical analysis. The distinct jump across the threshold is very intuitive for audiences. Both the actual and the predicted outcome can be illustrated on the basis of running variables. The graph will also give us the implication given the unexpected jump over the running variables. But, the selection of the bandwidth could affect the regression results and the graphical intuition. There is no unified rule to adopt the bandwidth. The cross-validation is suggested (Lee and Lemieux, 2009).[39]

The RD is a powerful design to capture the causal relationship with minimum assumptions. However, several assumptions are still needed. One of the key assumptions is that the observations are randomly distributed across the threshold. The distance of the observation will decrease the randomness. So our wish is to select the closest observations around the threshold. However, due to

[39] If the bandwidth is too narrow, then the results would be imprecise. If the bandwidth is too wide, the results would be less informative to tell the story.

the limit of data, we have to balance the numbers of observations and the risks of non-randomness. Due to this dilemma, Angrist and Rokkanen (2015) proposed a method to re-weight the running function in order to increase the observations without increasing the potential bias.[40] More recently, Dong (2016) argues that the result might be biased due to the kink effect and she proposed a general method, called "regression probability jump and kink (RPJK) design". Another key assumption is that the running variables are continuous (Hahn, et al, 2001). Moreover, normally RD model is to apply polynomial forms to estimate the effects of the running variables are linear.[41] This could introduce bias if the parametric forms don't fit the observations. Non-parametric models are proposed to mimic the running variables. Imbens and Lemieux (2008) argue that the standard kernel estimation is not appropriate for RD since we are interested in the effect of a boundary bound.[42] They suggest that local linear regression, series regression or sieve method would improve the results.

2.2.5 Sensitivity Test.

Although the applied economists have put many efforts into fitting the results,

[40] In their paper, they choose Propensity Score to estimate the running function. The CIA (Conditional independence assumption) and common support guarantee the function of the RD strategy.
[41] A practical guide for the polynomial terms is to represent every set of the polynomial terms to show the sensitivity of the results.
[42] The tradeoff between precision and bias is a fundamental problem of kernel regressions.

but in most of the work the results can only half explain by the regressions. In order to claim the results are robust, we need to rely on some assumptions, such as independence between error term and treatment variable. Therefore, various robustness and sensitivity tests are used to examine the assumption. Altonji et al (2005) propose a method, known as "AET", in which they examine if the results will vary when changing the correlation between the error term in the selection function and the error term in the reduced form. In the most recent developments, Oster (2016) propose another method in order to relax the two strict assumptions of AET, R-squared equals to one and equivalent importance between observed factors and unobserved factors.

2.3 The Returns to Education.

The return to education is one of the central topics in education economics. There are several major concerns in the empirical studies of return to education. Most empirical papers are based on the "Mincer Equation". First of all, the decision of education is highly endogenous, leading to stubborn selection bias into the results. Various studies tackle this problem with richer data, such as including the family background or early testing scores (Dearden, Lorraine, 1999; Meghir and Palme, 2005; Vignoles et al, 2010; Castex and Dechter, 2014). But using the family backgrounds to remove the ability bias has been criticized due to the fact that the family backgrounds are endogenous and don't strictly meet the validity assumption in the IV strategy.[43] Moreover, the results are also accompanied by the age effect and the cohort effect, given the fact that the returns to education are heterogeneous in terms of age and birth cohort (Mcintosh, 2006; Bhuller and Salvanes, 2017).[44] Third, the cognitive skills play important roles in determining both the decisions and the returns to education. Early scores have been used to proxy the personal ability in many studies. Many of them have argued that the cognitive skills have a strong positive correlation with returns. However, those skills might be multi-dimensional. The traditional cognitive skills may play different role compared with mechanical skills. Prada and Urzua (2017) argue that the mechanical ability is positively correlated with the returns but negatively correlated with the probability of university attendance. Fourth, the returns to education could be highly heterogeneous since there are many types of education targeting different groups of people. People who choose different education tend to have a rather diverse background, leading to complicated results.

Two trends are popular in dealing with the endogeneity when identifying the returns to education. The first trend is to use a comprehensive background to

[43] Trostel, Walker and Woolley (2002) examine the returns to education across countries with and without the control of family backgrounds. They point out that the traditional OLS results are underestimated and they also argue that the use of family backgrounds as proxies for children's ability is potentially problematic and an empirical question in the end.
[44] Bhuller and Salvanes (2017) examine the life cycle earnings on the basis of three identification strategies given a rich panel data in Norway. They find that the internal rate of return is around 11% and the standard Mincer equation tends to underestimate the return to the additional education.

control for the difference before attending to the education. The second trend is to use education reform as an instrument to estimate the effect of the additional education on returns, such as increasing compulsory schooling (Devereux and Hart, 2010), changing education path (Malamud and Pop-Eleches, 2010), the accessibility of schooling (Angrist and Krueger, 1991). The exogenous changes in years of education may help to eliminate the self-selection problem when estimating the returns.

Even though the results are not biased as a result of selection bias, the returns still could be highly heterogeneous as a result of the complex nature of returns to education. Many econometric tools have been proposed to tackle with regarding the heterogeneity among the students. Instrument strategy is a popular tool to estimate the returns to education as well. Given the study of Imbens and Angrist (1994), the LATE can evaluate the returns of people who change education choices as a result of the change in the instruments. However, in the context of heterogeneity, the people changed by the instruments might not be the same as the people who make the same decisions given the unchanged policy. Put it differently, the returns could be highly heterogeneous. Carneiro et al (2011) propose to estimate the returns to education by applying MTE given the same instruments. The marginal and average returns are not essentially the same, but the conventional average returns can be constructed on the basis of the marginal returns.

Chapter 3

Wage Return of Vocational Education: Evidence from the UK

3.1 Introduction： motivation, research questions and contributions.

A large amount of literature focuses on the issue of "Over-education" proposed by (Richard B. Freeman, 1976). 38% of university graduates are over-educated in their first job in the UK (Dolton and Vignoles 2002). The differences in the returns between academic and vocational education have attracted much attention from researchers. Compared to academic education, vocational education tends to provide firm and industry-specific skills. It not only prepares young students with specific skills but also offers a training route to enhance skills for academically oriented students.

In this chapter, I examine the crude differences in returns between vocational and academic education over birth cohorts. The results show that vocational qualifications have lower earnings compared to academic qualifications at similar levels and there are differences between different types of vocational education.

The main contribution of this chapter is to examine the effect of the education expansion on the returns to vocational education since a growing literature examines the effect of education expansion in the UK after 1988 on academic students (Walker and Zhu, 2008; Devereux and Fan, 2011) and finds no significant effects on the returns to university graduates, but little research has been done in terms of the impacts on vocational education. The increasing university graduates will compete with individuals with higher levels of vocational qualification directly, leading to an increasing pressure on vocational students. My aim is to test the stability of the returns to vocational qualifications. Although the returns to university graduates will also be changed during the education expansion, I examine the relative returns to vocational education compared to academic education in this chapter, rather than the causal effect of increasing university graduates on the returns to vocational students. If vocational students have unique advantages in the work-field, their returns will not be strongly affected by the increasing supply of university graduates. The results suggest that the wage caused by the education reform varies depending on the types of vocational qualification. I limit the range of age into 32-43 years old as a result of common support. It can allow me to examine the full productivity of individuals when they are in the middle of the career (Blundell et al, 2000).

However, unlike academic education, students in vocational education could be less homogenous in relation to the educational background, types of knowledge and so on (Billett 2011).[45] There are a wide range of types of vocational courses

[45] Different programs serve different purposes. For instance, the National Vocational Qualifications (NVQ) qualifications are more work-based and the Business and Technology Education Council (BTEC) qualifications are more course-based. Both

and providers. The structures of vocational education are more flexible. Individuals have more choices in relation to the training they need according to their plans. That encourages workers who are already in work to obtain more education through a vocational track. In my sample, around 70% of vocational students acquire their highest qualification with several years of working experiences and they can be classified as “returning students”. Others who obtain their highest qualification without working experiences can be defined as “fresh graduates”.[46]

From econometric perspectives, there are three difficulties in examining the difference in returns between the academic and the vocational education. The main difficulty lies in dealing with the differences in innate abilities between vocational and academic students before participating into the programs. Family background and early testing scores are commonly used to control for the innate personal ability (Dearden et al, 2002; Gavon, Conlon, 2005). Dearden et al (2002) argue that the effect of omitting innate ability would be mixed.[47] However, personal abilities would be very complex in this case due to the different nature

---

diversity of the program and limit of data may bring heterogeneity into the analysis.

[46] Vocational education can serve different purposes. Firstly, it may help individuals with their specific competence or skills related to the work. Secondly, vocational education can continuously help individuals with their skills and competence in the long term. Thirdly, vocational education can also help with the transition from one job to another, regardless of the transition is voluntary or forced. In the words, vocational education can help people with identifying the interest at the beginning, assisting with the career development later on and refining the specific competence and skills in the whole period of career (Billett 2011).

[47] After controlling family background based on National Child Development Study (NCDS), the results are similar to the results in Labour Force Survey.

of the two types of education.[48] They may have comparative advantages in different types of job.[49] Moreover, there are also difficulties in terms of examining the returns to education. The return to education often varies with age, birth cohort and level of education. McIntosh (2006) argues that the returns to vocational education remain constant throughout their working life. It is ambiguous to pin down the causal relation in terms of the difference in the returns between academic and vocational education on the basis of the early scores or the family background due to its strong magnitude of self-selection. A more convincing method is to take advantage of a quasi-experiment which induces people to switch between vocational and academic tracks.

The second issue is due to the various types of vocational education, leading to rather heterogeneous results. In the UK, there are several different vocational paths for students, such as National Vocational Qualifications (NVQs), Business and Technology Education Council (BTEC) etc. Moreover, students may change the path as a result of the education expansion. The third issue is due to the complex background of the vocational students compared to the academic students. There is a lack of information regarding their educational background.

---

[48] For instance, advanced cognitive skills may be distinct from communication skills and it is uncertain to tell which type of skill is more advanced or difficult to acquire since individuals are significantly different. Evidence has shown that the innate abilities have multiple dimensions (Prada and Urzua, 2017).

[49] From employer's perspectives, after taking into account the retraining cost between two types of workers, employers' hiring decisions will reflect the relative advantages and the degree of substitutability between the two types of education. Put it in another way, there is a tradeoff between higher learning ability with poorly matched knowledge and lower learning ability with precisely matched knowledge.

A significant proportion of vocational students obtain their highest qualifications with working experiences. The previous educational background might also be rather heterogeneous. So, I examine the differences in the returns between academic and vocational education on the basis of a quasi-experiment in the UK, trying to address the self-selection problem.

The rest of this chapter is organized as follows. Section 3.2 presents a brief literature review. Section 3.3 outlines my estimation methodology. Section 3.4 describes data and variables used in this chapter. Section 3.5 presents the results. Section 3.6 presents the conclusions and limitations.

3.2 Literature review.

The differences in the returns between academic and vocational education have been examined in many countries, but so far there are no unified conclusions.

A common empirical model to examine the difference in the returns is to add a dummy to represent the vocational worker in a regression. Hotchkiss (1993) examines the return to secondary vocational training on a training-related job after high school within two years. He finds that there is a wage premium for the vocational students in a training-related job. Neuman and Ziderman (1999) replicate his work based on the Israeli census data. They argue that Hotchkiss's work may be biased, since he focuses on young workers at the beginning of their career, leading to biased results. Vocational education may have advantages in work-related job, but the job might be at lower skill levels. It turns out that the average return is still lower than academic education. In the UK, Dearden et al (2002) find that the academic education tends to have higher returns than the vocational education based on the 1998 Quarterly Labour Force Survey (QLFS) and the 1991 sweep of National Child Development Study (NCDS). They also examine the selection bias by comparing the differences between QLFS and NCDS. After removing the selecting bias, the NCDS results are similar to the QLFS results without controls for the backgrounds.

On the other hand, considerable evidence suggests that vocational education may have a wage premium. Hawley (2003) shows that vocational education has higher returns for both men and women at the secondary and post-secondary level in Thailand. However, due to the fact that vocational students may have different innate abilities, the results might be strongly biased due to the omitted variable problem. Family background and personal ability indices are popular tools to tackle the selection bias. Moenjak and Worswick (2003) conclude that individuals from a well-do family are more likely to choose vocational education. After accounting for the family background, they find that the return to a vocational education is higher than a general education. Conlon (2005) finds that personal ability and family backgrounds have mixed impacts on choosing the types of education.[50] Meer (2007) builds a first stage model to remove selection problem with student's background and school quality. He finds that student from academic route may earn more by using the National Education Longitudinal Survey (NELS) of 1988.[51] Backes-Gellner and Geel (2014) find a positive return to a vocational education compared to an academic education.[52] They find that the two types of workers have a similar unemployment risk, but there is a higher return to vocational students in early career period. Vocational students have a lower risk of unemployment but the higher return disappears in the long-term. But different occupations and industries tend to have different results. Applying the Two-Stage-Least-Squares (2SLS) based on a family background to eliminate innate ability bias is very popular in the literature. However, it doesn't meet the exclusion restriction since a family background is not strictly exogenous in the wage equation.[53] Andersson, Nabavi and Wilhelmsson (2014) find that the vocational students tend to have 3-8 percent higher return than the students in an academic education after controlling for personal ability and family background based on Swedish dataset.[54] Polidano

---

[50] He argues that people are not willing to obtain the academic education although there is a wage premium compared to vocational education, because there is credit constraint among academic education. However, the determinants may change over time, especially with the development of the vocational education system in the UK.

[51] He divides the sample into general, academic, technical and business tracks under National Assessment of Educational Progress (NAEP) track definitions. The last two are vocational education.

The student's background includes information regarding their history of attending vocational classes, the scores of courses and parental information, reported by administrators.

[52] They use parental education as an instrument to solve the endogeneity.

[53] See Angrist and Pischke (2009).

[54] They have used various methods to examine the differences, such as IV,

and Tabasso (2014) argue that the returns may depend on the type of the vocational training programs in Australia. Their results suggest that the classroom-based vocational education with workplace specific training may achieve higher school completion rates and employment transition after controlling the pre-training academic scores. Due to the multi-dimensional nature of the personal abilities, only by controlling the innate abilities in the same way could not capture the impacts of the abilities in terms of affecting the outcomes.

The selection problem in relation to the study of the differences in the returns could be influenced by many factors. The simple OLS leaves the results hard to explain. Malamud and Pop-Eleches (2010) examine the effect of participating the vocational education on the returns on the basis of a natural experiment, launched in Romania. They argue that vocational students tend to have a lower employment rate and lower earnings compared to a general education. They argue that there are zero returns to the additional general education.[55]

Based on a rich dataset, a growing literature has already examined the differences in employment outcomes. However, since the education selection is rather endogenous, parental information and early testing score may not meet the strict exclusion restriction. With much less informative data regarding the

---

Hausman-Taylor estimators, fixed effect estimates and propensity score matching. Early scores are used as an approximation for ability. Parental information is used as IV.

[55] In order to address the selection bias, they use a natural education reform in Romania which pushed the vocational students into additional two years general education.

history of individuals, I examine the differences in returns on the basis of a quasi-experiment which was conducted in 1988 in the UK.

3.3 Empirical strategy.

There are no unified models for estimating the returns to vocational education. Some empirical papers investigated return to vocational qualifications based on the equation given the work of Hotchkiss (1993):

$$\boldsymbol{rlnwage_i = x_i \gamma + voc_i \beta + \varepsilon_i} \tag{1}$$

where "$\boldsymbol{rlnwage_i}$" is individual's log real wage. "$\boldsymbol{x_i}$" stands for individual and market's characteristics. The "$\boldsymbol{voc_i}$" equals one if one's highest qualification is a vocational qualification and that is also the variable of interest.[56] In this case, it stands for the total difference in returns between vocational education and academic education.

In order to examine the stability and substitutability of the returns to vocational education, I examine the effect of increasing supply of university graduates on a vocational education by applying DID. I address the selection problem using the education expansion after 1988. There was a substantial increase in university enrollment after the reform was enacted (Walker and Zhu, 2008). The education expansion pushed many polytechnics and higher education colleges into

[56] The 'voc' is defined as a dummy of vocationally related job or vocationally oriented program (Neuman and Ziderman 1999, Hotchkiss 1993).

becoming universities. It also relaxed the requirements for the new students. Those born after 1970 would have more opportunities to attend universities since they would be 18 years-old by 1988. On the other hand, the education expansion should not directly affect vocational education. Moreover, individuals who have vocational qualifications with working experiences may also become university graduates as a result of the reform. From my sample, the numbers of vocational students remain relatively constant between the 1965-1980 birth cohorts except for NVQs. The number of NVQs graduates increased significantly especially born after 1975.

Compared to the previous results which include both academic and vocational students, I only include observations with vocational qualification as their highest qualification in the DID strategy. The merit of this construction is that we can focus on the impacts on the vocational education. Although only the academic students experienced the education expansion directly, the vocational students may also be affected as a result of the increasing university graduates. However, they may also face more risks of spillover effect and Skill Biased Technology Change (SBTC). Without a richer data, I limit the birth cohorts, trying to minimize the biases as a result of the other policies. Moreover, I also manipulate the treatment group in order to test the robustness of the results. There are numerous types of vocational qualification and it is always changing over time, making it difficult to accurately categorize the vocational qualifications. Due to the limit of the data, I mainly examine the impacts on NVQ and BTEC.[57] The treatment group consists of higher levels of vocational qualification, including NVQ 5, NVQ4, NVQ3, BTEC higher, and BTEC national in my sample. The NVQ 3 and BTEC national belong to Level 3 which are similar to A-levels. Robustness checks have been done on the basis of the data without these marginal qualifications.

The supply of university graduates increased significantly and varied over time arising from the education reform after 1988. The enforcement of the reform is complex and took years. One concern is that there are more returning students who went to universities after 1976. That may bias the results if it is not controlled. Moreover, the results are also biased due to the compositional change in the innate abilities. Without a comprehensive dataset, I narrow the length of the birth cohorts from 1965 to 1979 in order to minimize the biases. I divide the whole expansion into three periods, pre-expansion, during-expansion and post-expansion (Devereux and Fan, 2011), 1965-1970, 1970-1975 and 1976-1980 respectively.

The hypothesis is that the increasing supply of graduates will affect the employment outcomes of higher levels of vocational student. Intuitively those new graduates will compete with higher levels of the vocational student directly

[57] Teaching and Nursing is a special class of vocational qualification. These qualifications are ranked higher in the data, but they are open to people with various education backgrounds, leading to rather ambiguous results. In this study, I don't differentiate them separately. Moreover, the data doesn't include higher levels of City & Guilds. In such case, there is no treated observations in the DID results.

and will decrease the returns to vocational education.

$$\boldsymbol{rlnwage_{it} = x_{it}\alpha + after_{it}\beta + hv_{it}\gamma + after_{it} \cdot hv_{it}\delta + \varepsilon_i} \qquad \textbf{(5)}$$

"$\boldsymbol{after_{it}}$" is denoted for individuals born after 1970. "$\boldsymbol{hv_{it}}$" denotes for higher vocational qualifications, including NVQ 5, NVQ4, NVQ3, BTEC higher, and BTEC national. "$\boldsymbol{x_{it}}$" is the vector of control variables. The coefficient $\boldsymbol{\delta}$ is our interest variable. It captures the effect of the education expansion on higher vocational qualifications. The algebra can be expressed as:

$$\delta = \left(\bar{y}_{H,2} - \bar{y}_{H,1}\right) - \left(\bar{y}_{L,2} - \bar{y}_{L,1}\right)$$

The first difference captures the effect of education expansion on the returns to individuals with higher vocational qualification. The second term captures common trend prior to the education expansion. δ will capture the effect of increasing university graduates on the returns to vocational education. The effect will be biased as a result of the compositional change between academic and vocational education and the personal ability bias from the education expansion.

There are several limitations based on the methodologies and dataset I used in this chapter. A potential problem is the composition of vocational and academic students may have changed as a result of the policy. Although the policy will not affect vocational education directly, it increases the opportunity for some vocational students to pursue a degree. Walker and Zhu (2008) don't find any fall in the premium for men and a large insignificant rise for female academic students. Moreover, the newly recruited graduates may have a different personal innate ability and educational backgrounds compared to the previous university graduates. I also apply quantile DID and narrow down the scope of samples to minimize the potential biases. Third, the spillover effect may exist in the results as a result of the definition of the qualifications. In the nature of education, it is hard to compare different types of education accurately. In this study, it might be controversial that NVQ 3 and BTEC national belong to the higher vocational qualification. More robustness checks are needed.

3.4 Data.

The data is drawn from 2001-2013 Quarterly Labour Force Survey (QLFS). The sample is restricted to individuals aged between 22 and 60 years-old who are male and employed. In this chapter, the academic route consists of GCSE, A-level, first degree and higher degree. Other qualifications fall within the scope of vocational education. The vocational qualifications are rather diverse. In the study, I focus on the typical types of vocational qualification such as NVQs, BTECs, and City & Guilds. In other words, I only select workers who report themselves holding a vocational qualification as their highest qualification. The dependent variable is the log of the real hourly wage. It is constructed as gross pay during the last pay period divided by the number of working hours of that pay period. In the DID results, I narrow down the age band into 32-43 years-old as a result of lack of common support and NVQs and BTECs since the data doesn't contain City & Guilds which is at the similar level as a first degree. Table 3.1 is the summary of variables.

<Table 3.1 Here>

<Table 3.2 Here>

Table 3.2 describes the comparison between types of qualification. The comparison is published by the UK government. It is used for the students and the employers to compare the qualifications. In QLFS, the variable doesn't have the option for BTEC professional.

<Figure 3.1 Here>

Figure 3.1 describes the proportions of the student with different qualifications given birth cohorts. It clearly shows that the university participation rate has a sharp increase for the cohorts born after 1970 and share of GCSEs was decreasing at the same time as a result of the education expansion. The higher education expansion pushed more GCSE graduates into A-level graduates and then the universities. However, compared to the previous university graduates, the newly recruited graduates may have lower personal abilities (Walker and Zhu, 2008). Moreover, the compositional change may not be limited to innate personal ability but also induce more diverse educational backgrounds. It gives people more opportunities to access the higher education. On the other hand, the

proportion of vocational education students remain relatively constant except for NVQs since the NVQs were introduced during that periods.

<Figure 3.2 Here>

Figure 3.2 describes the proportions of qualifications in different occupations between 2001-2003 and 2011-2013. The proportion of NVQs increases significantly in the ten years compared to other types of vocational education. However, the proportion of academic education increases in general and the vocational education has lower proportions in the labour market.

3.5 Empirical Results.

3.5.1 OLS results.

<Figure 3.3-3.5 Here>

Figures 3.3 to 3.5 show the big picture of how the returns to vocational qualification change over the birth cohorts. The negative effects are accompanied by time effects and age effects. It describes the differences in returns between academic and vocational education over birth cohorts from 1960-1990. The returns to vocational qualifications are lower than the academic qualifications in general. But the negative effects decreased significantly with the birth cohorts except for BTEC. It is interesting to notice that the number of NVQs graduates

also increased in the labour market. On the other hand, the returns of BTEC fell since 1965. The decreasing returns might be due to the education expansion. Increasingly vocational students went to pursue a university degree. Moreover, the selection bias may also reduce the returns since those students who have better educational backgrounds and higher potential earnings may pursue a degree as a result of the reform, leading to the decreasing returns for the vocational students. Moreover, since the average education increased rapidly during this time period, increasing students obtained degree or more A-levels. The returns are quite complex during the periods since the structure of the education system changed and the demand for the skills too. However, the results show that the returns to vocational education increased in general.

<Table 3.3 Here>

Table 3.3 describes the returns to various types of vocational qualification by OLS. It includes dummies to identify the returns to NVQ, BTEC, City & Guilds, and other vocational qualifications compared to academic qualification at the same nominal level. The reason for selecting these three vocational qualifications is that they have higher proportions among numerous vocational qualifications. The first column describes the returns in the whole sample. The results suggest that the vocational education tends to have lower returns compared to academic education in general. The second column is the results on the basis of 25-34 years-old between 2001 and 2003 QLFS. The third and fourth columns contain

2011-2013 and the sub-sample splits into 25-34 years-old and 35-45 years-old respectively. One of the merits of this arrangement is that the second column and the fourth column consist of a pseudo-panel and allow us to follow one cohort through their career after 10 years. In another word, the observations in the fourth column are the same birth cohort as the observations in the second column. The results suggest that the return to BTEC reduces significantly compared to other vocational qualifications. Although they have the lowest negative impact on vocational qualifications, they have the quickest decreasing rate in the middle of the career. Compared the third column with the second column, it shows that the returns to vocational qualifications decrease in general except for NVQs. That might be due to the fact that the NVQs were rapidly developing during that periods.

3.5.2 DID results.

<Table 3.4 Here>

The first column includes all sample between 1965 and 1980. The second column contains sample between 1965 and 1975. The third column contains sample between 1965-1969 and 1976-1980. The results suggest that the education expansion of university graduates reduce the returns to vocational education. The negative effects are larger and more significant in the post-expansion. The policies came out and became effective over time. There were more university graduates in the post-expansion. The simple comparison between the academic and the vocational education can be significantly biased due to the selection problem. The vocational students may have rather different innate abilities compared to academic students. In order to have a glimpse of the impacts of unobservable factors, I apply quantile regression with the DID strategy.

<Table 3.5 Here>

The results of quantile regression suggest that the unobservable factors play an important role in explaining the returns to qualifications. The returns to the higher vocational education increase with the quantile. The return to BTEC in the highest quantile is twice larger than the return on the lowest part. There no significant negative effects for the NVQs.

<Table 3.6 Here>

Table 3.6 shows the robustness check for the DID strategy. Compared to the previous results, I exclude the observations with NVQ 3 and BTEC national as their highest qualification since those qualifications are the marginal qualifications compared to a first degree. By excluding them we may examine whether the estimates change. In the results, both the magnitude and the significance of the results don't change. The effects of the education expansion on

vocational education are stable.

<Figure 3.6 Here>

Figure 3.6 shows how the negative effects of holding vocational qualification over the birth cohorts. The results show that the negative effect becomes significant born after 1974. One explanation is that there were many polytechnics and higher education colleges becoming universities after the “Further and Higher Education Act 1992”. It pushed many graduates to become university graduates directly. The university graduates increased significantly after 1992 which is also the time when people who born in 1974 become eligible for higher education.

3.6 Conclusion and limitations.

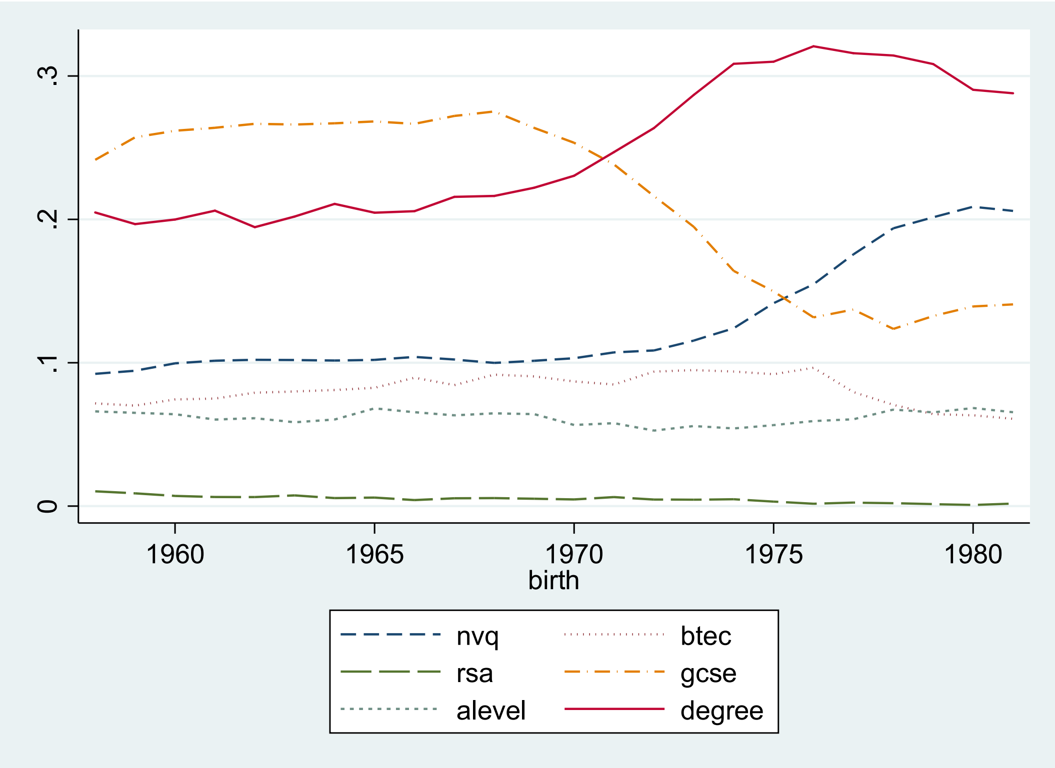
In this chapter, I first examine the crude differences between different types of vocational education and academic education over time. The return to vocational education exists over time compared to the academic education. The negative effects also vary by the types of vocational education. The return to BTECs decreases along with the education reform. The returns to NVQs and City & Guilds increase continuously with the increasing numbers of NVQ qualifications. Although the return to vocational education might be biased due to the age effect and innate abilities, it describes the relative differences in returns between types of education over birth cohorts. Moreover, the vocational education tends to have less promising career development in the future and the returns to vocational education become smaller in recent years.

Second, the results suggest the education expansion has negative effects on the returns to vocational education. The negative effects are smaller and less significant in the during-expansion periods when there are fewer graduates compared to post-expansion periods, suggesting that the quantitative effect of increasing graduates reduce the returns to vocational education.

In general, the results suggest that vocational education leads to lower earning compared to academic education in the middle of a career. The differences come from different educational backgrounds, personal abilities. The DID results suggest that the returns to vocational education are still affected by the over-supply of the university graduates. This chapter complements the literature in terms of the effects of educational expansion on vocational education. It implies that the overall effect of increasing university graduates on the society might be underestimated if it only counts the negative effects on university graduates.

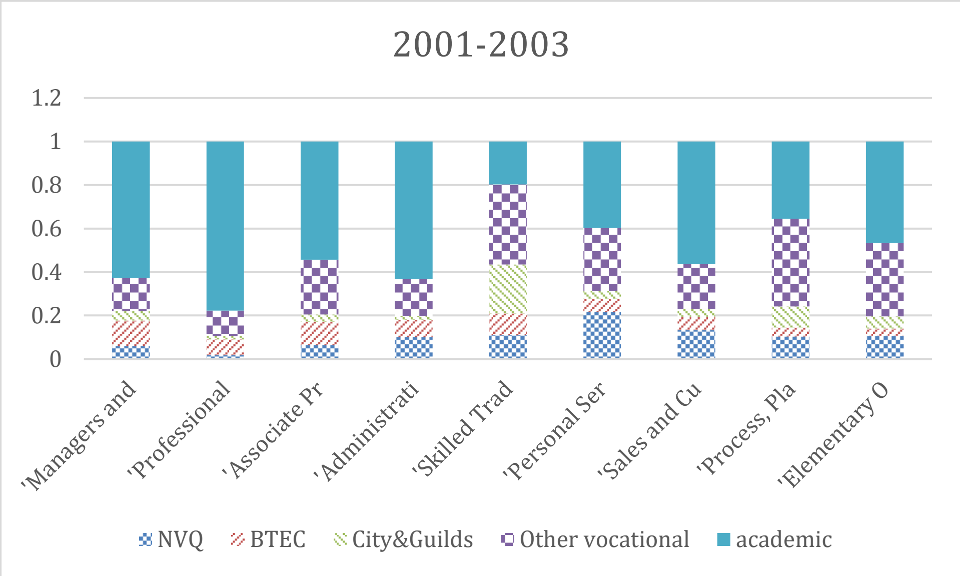
Figures and Tables:

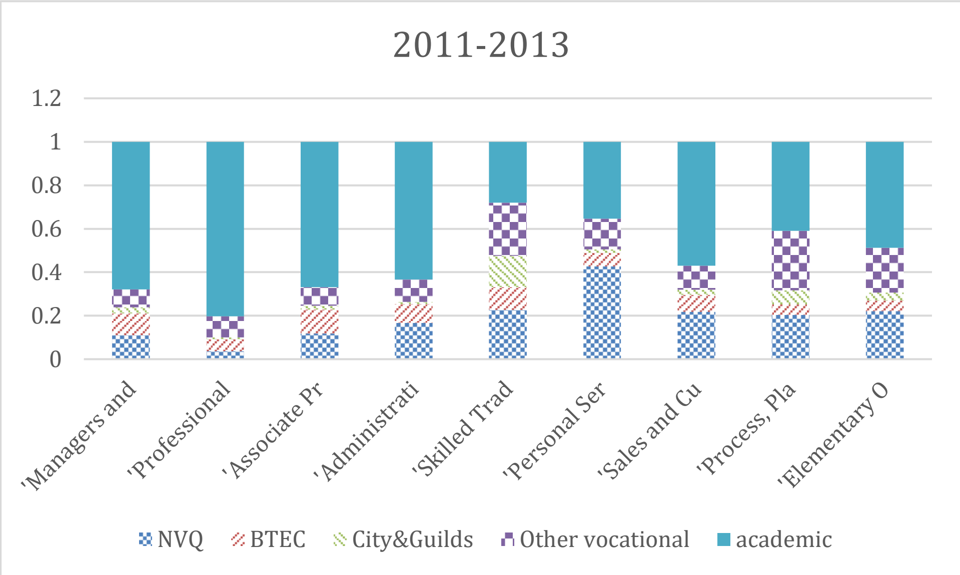
Figure 3.1 Proportion of different qualifications



Notes: Figure 3.1 includes all samples. It indicates the share of each qualification among the whole sample.

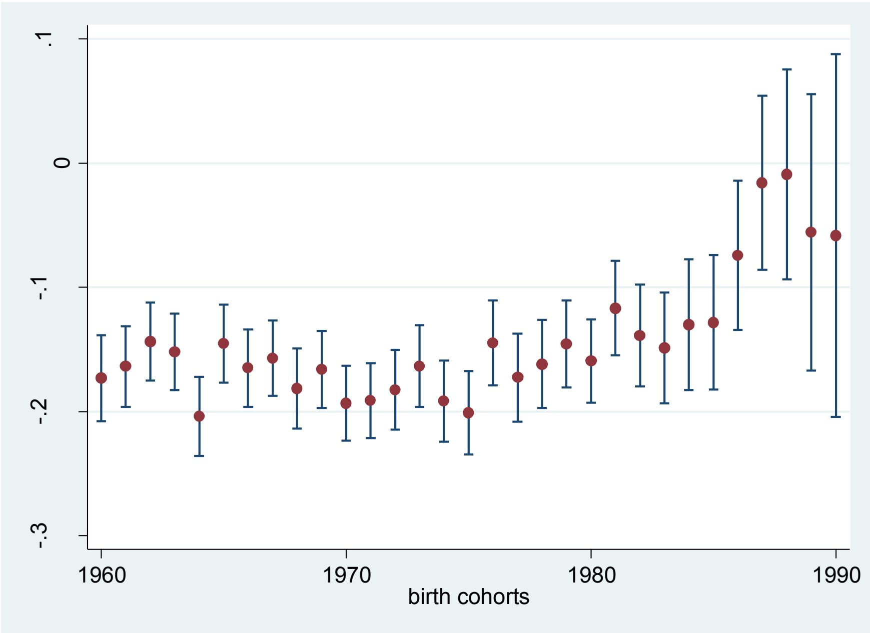
Figure 3.2 Vocational education in the workplace.





Notes: Each bar represents the percentage of vocational qualifications.

Figure 3.3 Return to other vocational qualifications over time



Notes: Control variables are education, experience, marriage, job training, disable, London, Full-time job, quarter, year, and industry. The control variables are the same in the following figures.

Figure 3.4 Return to NVQ over time

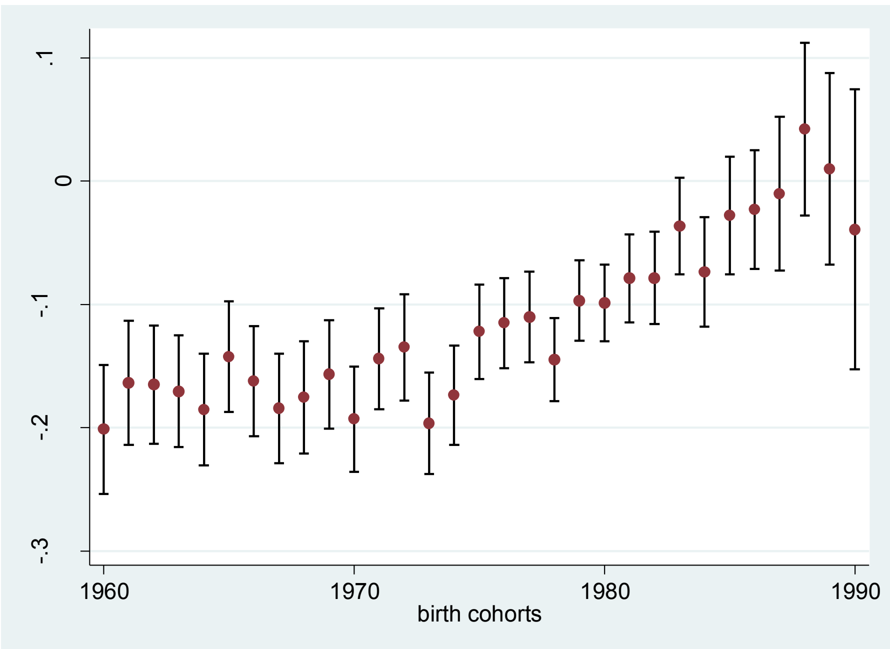


Figure 3.5 Return to BTEC over time

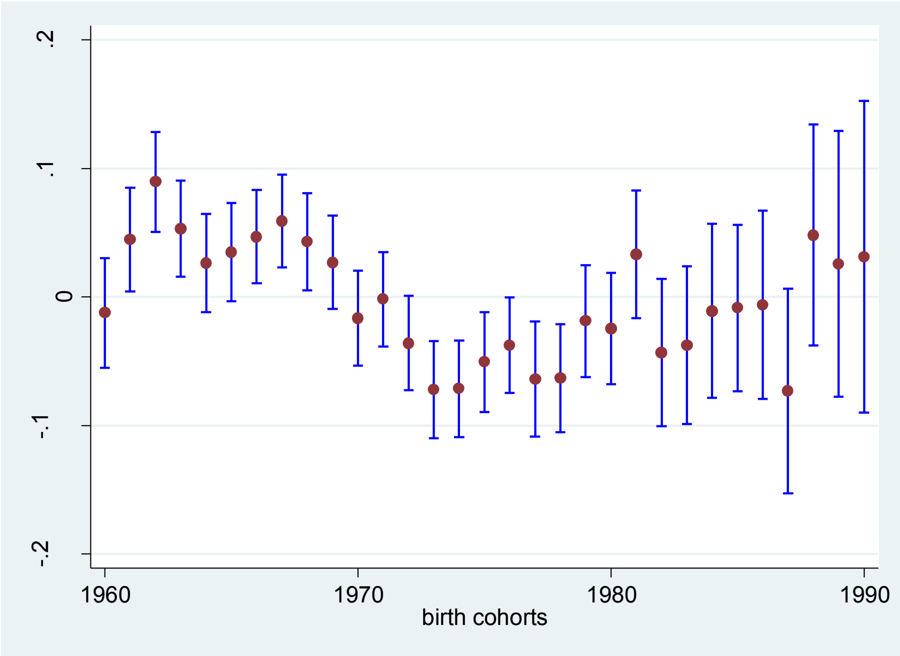
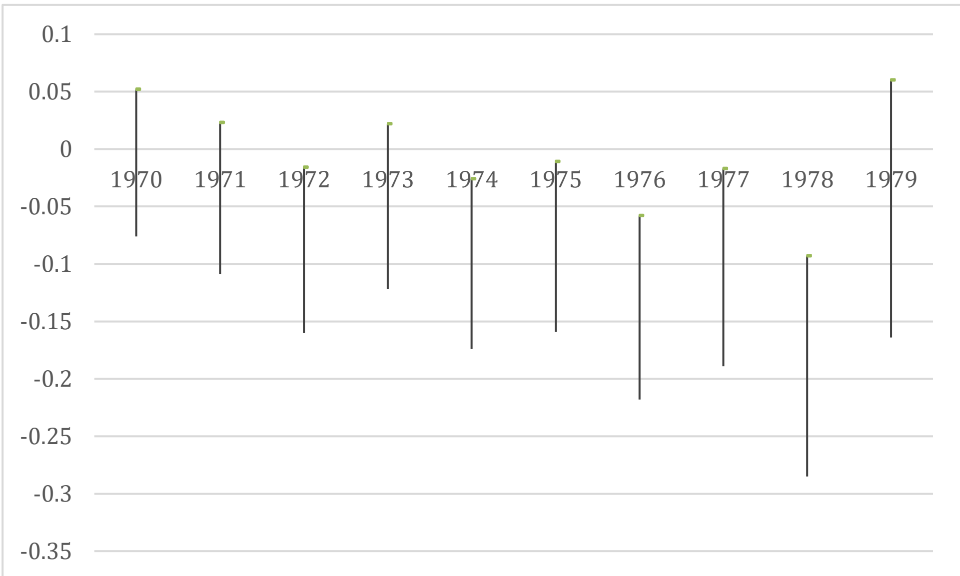


Figure 3.6 Returns to vocational education over birth cohorts



Notes: Each par represents the coefficient with the confidence interval.

Table 3.1 Summary statistics

VARIABLES	Explanation	Mean	sd
Age	Age of individuals	40.76	10.21
Firm50	If work in 0-49 company	0.436	0.499
Firm500	If work in 50-499 company	0.376	0.484
Firm1000	If work in 500-1000 company	0.188	0.391
Marr	If in marriage	0.619	0.486
Edu	Years of education	18.08	2.755
Exp	Work experience	22.68	11.11
Disable	If individuals are disable	0.129	0.335
Jobtrain	If receiving job related training	0.330	0.470
Full	Full time or part time job	0.773	0.420
London	If living in London	0.432	0.495
Permanent	Permanent job	0.957	0.204
rlnwage	Log of real wage	2.565	0.553
Voc	Vocational qualification	0.454	0.498
Nvq	Dummy for NVQ	0.119	0.324
Btec	Dummy for BTEC	0.075	0.264
Cityguild	Dummy for City&Guilds	0.037	0.189
Teaching	Dummy for teaching&nursing	0.034	0.182
Mature	Dummy for returning student	0.459	0.498
	Obs: 481482		

Notes: Edu is individual's education. Exp is experience. Birth is the year of birth. Disable, jobtrain, tenure, full, place, marr permanent, voc, and moti are dummies to identify if individuals have a disability, received job training from employers, have a full-time job, live in London, are married, have a permanent job, hold a vocational qualification and received their highest qualification with work experience respectively. Firm50, firm500, firm1000 are dummies to represent the scale of companies. The DID results are based on observation's age from 32-43 years old since observations are only over-lapped in this age band.

Table 3.2 Qualification comparison

NVQ level	NVQ	BTEC	Equivalent (Academic) Qualification
1	NVQ 1	BTEC 1	GCSE (D-G)
2	NVQ 2	BTEC 2	GCSE (A-C), City & Guilds 2
3	NVQ 3	BTEC professional award, certificate and diploma level 3	AS / A level, City & Guilds 3
4	NVQ 4	BTEC professional award, certificate and diploma level 4, 5, 6	HNC/HND, Diploma, Bachelor's degree, City & Guilds Licentiateship, Associateship, and graduateship
5	NVQ 5	BTEC professional award, certificate and diploma level 7	Master's degree, City & Guilds Membership and Fellowship

Notes: See

https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels

Table 3.3 Returns to vocational education.

VARIABLES	QLFS 2001-2013 All sample	QLFS 2001-2003 Aged 25-34	QLFS 2011-2013 Aged 25-34	QLFS 2011-2013 Aged 35-45
nvq	-0.202***	-0.147***	-0.128***	-0.244***
	(0.004)	(0.013)	(0.012)	(0.014)
btec	-0.019***	-0.027**	-0.044***	-0.046***
	(0.004)	(0.011)	(0.017)	(0.013)
cityguild	-0.165***	-0.143***	-0.102***	-0.220***
	(0.005)	(0.015)	(0.032)	(0.021)
othervoc	-0.179***	-0.158***	-0.174***	-0.206***
	(0.003)	(0.010)	(0.016)	(0.014)
edu	0.066***	0.086***	0.069***	0.064***
	(0.000)	(0.002)	(0.002)	(0.002)
exp	0.041***	0.087***	0.069***	0.076***
	(0.000)	(0.004)	(0.005)	(0.008)
expsqua	-0.001***	-0.002***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
moti	0.088***	0.088***	0.043***	0.065***
	(0.002)	(0.008)	(0.010)	(0.009)
marr	0.111***	0.086***	0.078***	0.115***
	(0.002)	(0.007)	(0.009)	(0.009)
jobtrain	0.084***	0.069***	0.067***	0.069***
	(0.002)	(0.007)	(0.009)	(0.009)
disable	-0.084***	-0.092***	-0.102***	-0.098***
	(0.003)	(0.012)	(0.017)	(0.014)
place	0.127***	0.159***	0.122***	0.124***
	(0.002)	(0.007)	(0.008)	(0.008)
full	0.310***	0.332***	0.337***	0.368***
	(0.005)	(0.021)	(0.018)	(0.021)
Constant	0.584***	-0.045	0.231***	0.069

	(0.012)	(0.060)	(0.065)	(0.110)
Observations	229,318	15,123	10,623	14,446
R-squared	0.293	0.304	0.282	0.278

Notes: Both vocational and academic qualifications are included in the OLS. Academic qualifications are the default variable in the regression. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.4 DID results

VARIABLES	(1) QLFS 1965-1980	(2) QLFS 1965-1975	(3) QLFS (1965-1969) & (1976-1980)
after	0.038*	0.044*	0.038
	(0.019)	(0.024)	(0.042)
hvnvq	0.254***	0.240***	0.251***
	(0.018)	(0.020)	(0.018)
hvbtec	0.418***	0.418***	0.411***
	(0.016)	(0.017)	(0.016)
afterhvnvq	-0.041*	-0.011	-0.066**
	(0.022)	(0.026)	(0.029)
afterhvbtec	-0.091***	-0.071***	-0.133***
	(0.019)	(0.023)	(0.028)
edu	0.016***	0.022***	0.018***
	(0.003)	(0.005)	(0.005)
exp	0.048***	0.052***	0.053***
	(0.008)	(0.011)	(0.012)
expsqua	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
moti	-0.015	-0.002	-0.023
	(0.011)	(0.014)	(0.014)
marr	0.087***	0.099***	0.102***
	(0.008)	(0.009)	(0.010)
jobtrain	0.048***	0.049***	0.054***
	(0.008)	(0.009)	(0.010)
disable	-0.088***	-0.079***	-0.069***
	(0.012)	(0.014)	(0.015)
place	0.131***	0.141***	0.131***
	(0.008)	(0.009)	(0.010)

full	0.366***	0.352***	0.353***
	(0.020)	(0.025)	(0.028)
Constant	1.125***	0.919***	1.073***
	(0.119)	(0.177)	(0.191)
Observations	13133	9496	7319
R-squared	0.253	0.240	0.278

Notes: In the DID strategy, only vocational qualifications are included. “hvnvq” stands for individual who hold a higher than NVQ 3 as highest qualification. “hvbtec” stands for individuals who hold a BTEC higher or a BTEC national. “afterhvnvq” and ”afterhvbtec” are the interaction terms and also the variable of interest. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.5 Quantile DID

VARIABLES	0.1	0.3	0.5	0.7	0.9
after	0.005	0.042*	0.020	0.018	0.027
	(0.028)	(0.022)	(0.021)	(0.022)	(0.030)
hvnvq	0.190***	0.238***	0.270***	0.272***	0.276***
	(0.026)	(0.021)	(0.020)	(0.021)	(0.028)
hvbtec	0.322***	0.408***	0.429***	0.429***	0.460***
	(0.023)	(0.018)	(0.017)	(0.018)	(0.024)
afterhvnvq	-0.015	-0.015	-0.020	-0.031	-0.042
	(0.032)	(0.025)	(0.024)	(0.025)	(0.034)
afterhvbtec	-0.056**	-0.089***	-0.071***	-0.065***	-0.107***
	(0.028)	(0.022)	(0.021)	(0.022)	(0.029)
edu	0.009**	0.015***	0.014***	0.021***	0.016***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
exp	0.056***	0.049***	0.049***	0.051***	0.038***
	(0.012)	(0.009)	(0.009)	(0.009)	(0.012)
expsqua	-0.001***	0.001***	-0.001***	0.001***	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
moti	0.041***	0.004	-0.022*	-0.033***	-0.058***
	(0.016)	(0.012)	(0.012)	(0.012)	(0.016)
marr	0.070***	0.093***	0.085***	0.086***	0.010***
	(0.011)	(0.009)	(0.008)	(0.009)	(0.011)
jobtrain	0.059***	0.051***	0.050***	0.044***	0.036***
	(0.011)	(0.009)	(0.009)	(0.009)	(0.012)
disable	-0.082***	-0.106***	-0.081***	-0.083***	-0.066***
	(0.017)	(0.013)	(0.013)	(0.013)	(0.018)
place	0.113***	0.124***	0.129***	0.135***	0.149***
	(0.011)	(0.009)	(0.008)	(0.009)	(0.011)
full	0.398***	0.390***	0.352***	0.347***	0.276***
	(0.030)	(0.023)	(0.022)	(0.023)	(0.031)
Constant	0.737***	0.892***	1.150***	1.198***	1.746***

	(0.174)	(0.135)	(0.131)	(0.136)	(0.181)
Observations	13,133	13,133	13,133	13,133	13,133
R-squared	0.132	0.164	0.169	0.163	0.155

Notes: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.6 DID robustness check

	(1)	(2)	(3)
VARIABLES	All sample	Pre-expansion	Post-expansion
after	0.0379*	0.0329	0.0451
	(0.020)	(0.025)	(0.045)
hvnvq	0.260***	0.247***	0.256***
	(0.018)	(0.020)	(0.018)
hvbtec	0.452***	0.451***	0.443***
	(0.016)	(0.018)	(0.017)
afterhvnvq	-0.045**	-0.014	-0.076**
	(0.022)	(0.026)	(0.030)
afterhvbtec	-0.094***	-0.073***	-0.140***
	(0.020)	(0.023)	(0.030)
edu	0.012***	0.015***	0.017***
	(0.003)	(0.005)	(0.005)
exp	0.046***	0.048***	0.051***
	(0.009)	(0.011)	(0.013)
expsqua	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
moti	-0.034***	-0.031**	-0.030*
	(0.012)	(0.016)	(0.017)
marr	0.088***	0.100***	0.105***
	(0.008)	(0.010)	(0.011)
jobtrain	0.051***	0.054***	0.058***
	(0.008)	(0.001)	(0.011)
disable	-0.079***	-0.072***	-0.050***
	(0.013)	(0.015)	(0.017)
place	0.130***	0.140***	0.126***
	(0.008)	(0.010)	(0.011)
full	0.349***	0.342***	0.356***
	(0.022)	(0.026)	(0.030)

Constant	1.234***	1.135***	1.103***
	(0.127)	(0.189)	(0.207)
Observations	10914	7875	6036
R-squared	0.278	0.267	0.306

Notes: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Chapter 4

Does age-dependent minimum wage affect employment? evidence from UK

4.1 Introduction.

The National Minimum Wage (NMW) in the UK was first introduced in 1999 with different rates for the age bands, 18-21 years-old and 22-years-old and above. The age-dependent minimum wage structure took its current form in 2004, when a lower rate of the minimum wage was introduced for 16-17 years old at than 18-21 years old band.[58] The age-dependent minimum wage is used to regulate the flow of young workers into the labour market. By differentiating the minimum wage it may give employers incentives to recruit younger workers in favour of older workers in order to minimize the cost. Therefore, it may help younger workers into employment since they are the most vulnerable in the labour market. In the standard economics textbook, higher minimum wage results in reduced employment rate in a perfectly competitive labour market, where the marginal labour cost equals the marginal product of labour. On the other hand, in a monopsonistic labour market, marginal labour cost will exceed the wage rate with an upward-sloping labour supply.

[58] In 2010, the adult minimum wage age cut-off changed from 22-years-old to 21-years-old.

4.1.1 Effects on the stock of labour.

Many studies have examined the effect of the introduction of the minimum wage on employment rate based on survey data or establishment level data (Card and Krueger, 1994; Machin et al, 2003; Stewart, 2004; Arulampalam et al, 2004; Dickens et al, 2015).[59] A consensus on minimum wage is that a modest increase in minimum wage will not lead to a large reduction in employment empirically, whereas, it tends to compress the wage distribution (Machin et al, 2003).[60]

One of the explanations for this disparity is that firms manage to reorganize the production process to offset the increasing minimum wage (Draca et al, 2011, Riley and Bondibene, 2017).[61] However, a major weakness of the study based on survey data is the lack of examination of the compositional changes within firms (Gioliano, 2013). Gioliano (2013) examines the impact of increasing minimum

[59] Stewart (2004) examines the effect of NMW on employment probability by using Difference-in-Difference (DID) based on the British Household Panel Survey (BHPS), the Quarterly Labour Force Survey (QLFS) and the New Earnings Surveys (NES) and concludes that there is no significantly negative effect on employment. While the results might be biased due to the spillover effect, Dickens and Manning (2004) argue that there is little evidence in terms of spillover effect.
Arulampalam, Booth and Bryan (2004) argue that there is little evidence that the introduction of the minimum wage will increase the job-related training.
Dickens, Riley and Wilkinson (2015) examine the employment effect on the most vulnerable group in the UK, namely part-time females, based on QLFS and NES by using DID. They conclude that the increase in minimum wage will decrease the employment probability for part-time females and it will get worse in the recession. In their work, they also show that the effect of minimum wage can be very diverse depending on the groups. However, the ambiguous effect of Education Maintenance Allowance (EMA) and new minimum wage still may bias the results.

[60] Card and Krueger (1994) examine the sign and magnitude of the effect of introducing minimum wage and explain the results in the context of monopsonistic power of firms.

[61] Riley and Bondibene (2017) examine the firm's reaction toward to the increasing minimum wage in the UK based on Financial Analysis Made Easy (FAME) data, which comprises more firms compared to Annual Survey of Hours and Earning (ASHE) (Draca et al, 2011). Following Draca et al (2011), they apply DID to firm-level data to separate the treatment group from control group on the basis of average labour costs.

wage on employment, especially focusing on teenagers. Given the unique personnel data from US retail firm, she concludes that the increase in relative minimum wage of teenagers raises teenager employment and the increase in average wage leads to a negative but statistically insignificant impact on overall employment. But many unaddressed questions on the effect of increasing minimum wage from different dimensions still remain (Metcalf 2008).

4.1.2 Effects on the flow of labour.

Another angle to study the impact of increasing minimum wage is to examine the flow of labour rather than the stock of labour. From recent literature, Brochu and Green (2013) argue that higher minimum wage is associated with lower hiring rate and lower job separation rate using Canadian data from 1979 to 2008. Unskilled workers are most likely affected by the increasing minimum wage.[62] Due to the limited numbers of provinces in Canada, the constructed province level variables can’t allow them to examine the bias arising from heterogeneity by which it is argued (Dube et al, 2016). Apart from the match quality model in Brochu and Green (2013), Dube et al (2016) apply the job-ladder model and

[62] Their study differs from previous literature on three aspects. First, they study the transition rates before and after the change rather than in the transition. Second, they focus on the new hires who have less than one year tenure. Third, they also examine the impact on unemployed and inactive observations. Based on the Mortensen-Pissarides search and matching model, they argue that firms are less willing to terminate the contract because the opportunity cost of new search is higher after increasing minimum wage due to the cost of screening.

argue that the dis-employment effect of teenagers can be significantly changed due to the heterogeneous trends in the US. Their results suggest that increasing minimum wage has strong negative effects on job separation rate, job accession rates, and turnover rate but not on the stock of labour. Kabatek (2015) examines job accession rate and job separation rate based on regression discontinuity and individual-level administrative data. He argues that individuals may be made redundant shortly before becoming eligible for higher minimum wage level.

Although the introduction of minimum wage and the flow of labour in the U.S. have been widely examined, we still know relatively little in terms of the effect on labour supply in the UK.

4.1.3 Heterogeneous effects.

To my knowledge, Dickens et al (2014) is the first the UK study that applies Regression Discontinuity (RD) to examine the effect of the increase in minimum wage on labour supply. Their results suggest that after becoming eligible for higher minimum wage rate, individuals have higher employment probability based on RD. Although their results are very robust, the increase in employment probability can't be entirely explained by the increase in labour supply since employers may still reorganize production shortly after introducing the minimum wage. Moreover, the effects could be heterogeneous for sub-groups.

After becoming eligible for higher minimum wage rate, individuals have a higher expected return, leading to higher search intensity and labour supply. According to the search and matching theory, it results in higher matching rate and productivity. The discontinuity also leads to more competition in the labour market when higher skilled workers and lower skilled workers increase labour supply at the same time. The exogenously increasing competition may induce a crowding out effect on lower-skilled workers when the labour market is tight (Dolado et al, 2000; Lene, 2011).[63] Martin and Pierrard (2014) propose a theoretical model and argue that on-the-job search will lead firms to open more vacancies and crowd out unemployed workers in the job search. The crowding out effect may also depend on the tightness of the local labour market (Addison et al, 2013).[64] The crowding out effect is still a controversial topic.[65] Gautier et al (2002) find very thin evidence that higher-skilled workers crowd out lower-skilled workers during the recession in the Netherlands, although they do not focus on minimum rate level job.

---

[63] Lene (2011) proposes a theoretical model to illustrate an increase in the supply of relatively high skilled workers will reduce employment opportunity for lower skilled workers.
Dolado et al (2000) extend the search model (Van Ours and Ridder 1995) to explain the crowding out effect in Spain. But their assumption regarding leisure is perhaps too strong, since individuals will wait until they find a skilled job.

[64] However, crowding out effect is not easy to observe directly. Comparing employment probability over qualifications is ambiguous when higher skilled workers may find a job not covered by minimum wage. Given initial Figures of employment rate, minimum wage will mostly affect individuals with qualification below A-levels. Addison et al (2013) suggest that individuals with minimum wage rate hit mostly in recession.

[65] The crowding out effect being examined by the stock of labour is ambiguous to some extent. It may be explained more precisely by examining the flow of labour, turnover rate, job separation rate, and job accession rate.

Except for the employment probability, the types of a job may also imply different employment strength. A temporary job can be seen as a "stepping-stone" for the later job and is associated with a lower wage, lower job satisfaction, and on-the-job training opportunities compared to a more permanent job (Booth et al, 2002; Engellandt and Riphahn, 2005).[66] After the increase in the expected return, individuals have more reasons to find a full-time or a permanent job (Card and Krueger, 1994; Nunez and Livanos, 2015).[67] Individuals may search for a job starting from the "good" job, leading to a queue in each vacancy and those good jobs are filled up firstly after the matching between employee and employer. So the full-time or permanent job may illustrate the relative strength between different types of workers in the labour market.[68]

4.1.4 The contribution.

Compared to the previous literature, I examine the effect on employment

---

[66] Engellandt and Riphahn (2005) argue that workers with a temporary contract may exert higher effort in order to transfer into a permanent contract under the hypothesis that employers will screen workers through temporary contract.

[67] Card and Krueger (1994) discuss the possible substitution between full time job and part time job due to both employee's and employer's motivation.

[68] Who are the compliers to this policy? Obviously, those who can find adult wage before becoming eligible for would not comply with this policy. They are the never-takers in terms of applied econometrics. Intuitively those whose reservation wages are lower than youth development minimum wage rate will always try to find a job even without the policy. So they are the always-takers. It is expected that those observations whose reservation wage are higher than development minimum wage rate but lower than adult minimum wage will step into labour market or have more desire to find a job. Falk et al (2006) show that there is a strong relation between minimum wage and reservation wages based on their unique laboratory experimental data. They argue that the increasing reservation wage doesn't decrease even after the temporarily increasing minimum wage falls.

probability in subgroups and highlight the possibility of a crowding out effect in the recession brought by higher competition after becoming eligible for higher minimum wage rate. The increasing minimum wage will increase the labour supply and also increase the competition between different groups, leading to crowding out effect.

Compared to Dickens et al (2014) which examine the effects on the whole, lower skilled group, I divide the lower skilled workers into subgroups on the basis of their qualifications and the results show that there exists heterogeneous effects given their qualifications after becoming eligible for the higher minimum wage rate. Since the samples are mainly from the "Great Recession" period, the probability of the crowding out effect is also highlighted in this chapter. Moreover, I also discuss more potential biases to the results and other potential employment effects which have been neglected in the previous literature. The main results suggest that there is no effect on the employment probability among individuals whose highest qualification is below GCSE. Among the GCSE group, the positive discontinuity comes from individuals with 5 or more GCSEs and there is no significant effect for individuals with less than 5 GCSEs. As a potential bias for RD results, lower skilled workers might be made redundant (Kabatek, 2015). I examine the discontinuity shortly before the age threshold and the results suggest that there are no replacement effects among employees in the UK. After disaggregating the results by male or female, the discontinuity focuses on the male. Moreover, individuals with higher numbers or better grades of GCSE tend to have a higher probability of finding a full-time job or permanent job after becoming eligible for higher minimum wage rate. Individuals with lower skills don't have any significant discontinuity. The results suggest that policy makers should take those possible non-negligible adverse effects into consideration.

The chapter is organized as follows. The methodology is discussed in section 4.2. Section 4.3 introduces the dataset. The results are presented and discussed in section 4.4. Section 4.5 concludes.

4.2 Regression discontinuity regression.

RD design is a quasi-experimental design in which the probability of receiving treatment will be changed discontinuously across the threshold as a function of control variables (Hahn et al 2001). It is an increasingly popular method in applied econometrics (Imbens and Lemieux 2008). However, in this study, one potential problem is that the distance is recorded as monthly rather than daily (discrete to some extent), leading to potential ambiguous bias into the results (Lee and Card, 2008).

In this chapter, I examine the age discrimination based on discontinuity around a well-defined age cut-off. As for the econometric model, I apply interaction between the discontinuity dummy and the distance to the cut-off point to allow for the slope changing after crossing the age threshold, using constant, linear, and quadratic models.[69] This econometric framework exploits the discrimination from employers based on sharp regression discontinuity model:

$$Y_{ia} = \alpha + \beta TREAT_{ia} + \delta(a) + \gamma X_i + \varepsilon_{ia} \quad (1)$$

where Y_{ia} is outcome variable for individual i of age a. X_i are the control variables. Probit model is used to evaluate the effects. Here TREAT is our interest variable, indicating the effect of discontinuity. It equals to one if an individual is eligible for the higher minimum wage. β is the effect of discontinuity induced by the increasing minimum wage. $\delta(a)$ is called time function which captures the effect of duration. A key assumption in RD strategy is that $\delta\ (a)$ is a continuous function. The samples before and after birthday are assumed to be randomly assigned. That guarantees the treatment variable is the only source of discontinuity.

$$\delta(a) = \theta_1 duration + \theta_2 duration^2 + \theta_3 duration * TREAT + \theta_4 (duration * TREAT)^2 \quad (2)$$

It is difficult to select time control function. Here robust tests in practice are necessary. By including control variables to minimize the error term we need to mimic the parametric form. Under the parametric framework, the estimates might be biased due to exogeneity. After the test of the validity of RD design, the control variables are used to decrease the variability. We can also decrease the

[69] Non-parametric smoothing is not applied here due to the limit of the data.

bias by narrowing down the window length to some extent, but the number of observations is decreasing as well. This is the tradeoff between precision and bias.

The essence of RD is to compare the treatment group on the right-hand side of the cut-off point and comparison group on the left side of the cut-off point. The control group from marginally below the threshold is valid counterfactual for the treatment group from marginally above the threshold (Hahn et al, 2001). Around the threshold, the treatment status is independent of all variables no matter observable and unobservable like a random assignment (Lee, 2008). Here the instruments are the distances to the month of birth. Put it differently, individuals have imprecise control of the treatment status. In this case, the randomness of the month of birth can guarantee the randomness of the treatment from the threshold.

This chapter focuses on the effect of labour supply when individuals become eligible for the higher minimum wage rate. In this chapter, I examine the effect of the increasing minimum wage on labour supply and the effect on finding different types of job for different demographic groups.

One assumption here is that there is no administrative cost for employer associated with firing and hiring, since the employer may prefer existing workers or workers who are already eligible for the higher minimum wage if the administrative is non-negligible. The discontinuity would be larger when this employer's preference exists. In this case, it is a rather empirical question and it is a reliable assumption since employer posts a job given their labour demand of production process and the workers across cut-off point would be identical to employers.

Another potential bias comes from the employer's motivation for replacement shortly before workers become eligible for higher minimum wage rate. It is ambiguous since employer's firing decision may not only depend on individual's age but also their own production process.[70] Kabatek (2015) argue the discontinuity might be the artificial results induced by replacement. In my chapter, I examine the replacement effect by running regressions on the periods shortly before the age threshold. The evidence suggests that a significant replacement effect is not found in the UK based on Quarterly Labour Force Survey (QLFS).[71] I will discuss this issue in details later and put the results into the appendix.

Moreover, individuals may change their labour supply behaviour before becoming eligible for the higher minimum wage rate. The anticipation effect will lead to biased results. Dickens et al (2014) intensively discuss the potential motivation behind anticipation behaviour and check the effect based on

---

[70] It is an empirical and complex issue and it may exist to some extent, although a more important issue is the magnitude of replacement.

[71] Moreover, there is no significant change from the perspective of macro-economy given the current data generating process. My whole samples are in a recessionary period. All of those I described above make sure that there is no change in demand side across the cut-off point.

Difference-in-Difference among 20 years old compared with 21 years old before and after the introduction of the policy. They argue that the effect is small when they become eligible for higher minimum wage rate.

Third, the policy regarding the probationary period before a permanent contract may also bring potential bias into the results since employers may have the motivation to recruit more workers when the hiring or firing cost is lower when workers are in probation. In other words, employers may take advantage of the probation rule to recruit more workers 6 months before the age threshold at 22th birthday under the assumption that employers have maximum 6 months to screen worker's full productivity. If probation effect exists, then it will upward bias the results. I examine this effect based on DID and don't find significant probation effect in the sample. The results are in the appendix2.

Lastly, the results might be also biased by time effect and kink effect (Fidrmuc and Tena, 2013; Dong, 2016).[72] The time effect mainly comes from the changing employment strength over time, leading to the failure of Conditional Independence assumption (CIA) of running variable.[73] Due to the limit of the data in this study, the simplest way to minimize the bias coming from time effect

---

[72] Dong (2016) proves that kink effect can also make a significantly artificial jump around the cut-off point.

[73] The age effect could be checked and corrected by applying inverse probability weighting and the conditional independence test could also be applied (Angrist and Rokkanen, 2015). Alternatively, augmented inverse probability weighting (AIPW) can be implemented as it includes all set of information both in the propensity score and the outcome function. Under the assumption that propensity score and outcome function are correctly estimated, AIPW would be unbiased for ATE (Glynn and Quinn, 2010). However, it does need a more comprehensive dataset to illustrate the demographic characteristics across the threshold.

is to narrow down the periods around the cut-off point and apply polynomial terms to mimic the pattern.

4.3 Data and statistics.

I pool five years QLFS together from 2008 to 2012. It contains month of birth which is used to calculate the distance from the month of the survey to the month of birth. Before Oct 2010, an increase in minimum wage is due on one's 22nd birthday, but the age threshold is changed into 21-years-old after 2010. The sample is restricted to individuals who are 21-years-old or 22-years old before 2010 or 20-years-old and 21-years-old after 2010. Observations are ranked on the basis of their distances between the month of the survey to the month of birth. There might be a non-negligible effect on employment probability of individuals with lower levels of education.[74] Therefore, instead of focusing on 18-years-old threshold, the effect of the increase in minimum wage on employment probability of individuals who are turning into 21-years-old is cleaner. In order to get rid of the measurement bias, I drop individuals whose month of birth equals to month of survey since it is unclear if individuals have passed their birthday.

[74] A-level graduates may also find a job in the range of minimum wage.

4.4 Results.

4.4.1 Employment probability.

Individuals with lower levels of education are expected to be more affected by the increasing minimum wage. Although individuals with more years of education may still find a job which is paid around the minimum wage, the employment probability of individuals with A-levels are not largely affected by minimum wage in my results. The main results focus on the whole sample regardless of their sex, mostly due to limited sample size. After splitting on male or female, the results are less robust. In order to balance between bias and efficiency, I examine the discontinuity in different periods and add different polynomial terms as a time function.

Figure 4.1 describes the employment rates given different qualifications on the basis of distances. The upper left figure is the employment rate of individuals with qualification lower than GCSE. Upper right and lower left are the employment rates of individuals with fewer than five GCSEs and five or more GCSEs, respectively. The last is for individuals whose highest qualification is A-level. From the simple pattern of employment rates, there is no clear evidence in terms of the relation between minimum wage and employment probability for individuals with any qualifications except for individuals with five or more GCSEs. They have a clearer increasing trend when they are eligible for higher minimum wage rate. Obviously, A-level students tend to have better chances to get a job compared with individuals with less education. Results for individuals with qualification lower than GCSE are rather ambiguous since it includes a variety of qualification. Putting it here is to show the continuity of qualifications.

<Figure 4.1 Here>

Table 4.1 shows the Probit regression results of individuals with qualification lower than GCSE and GCSEs. The parameter of "TREAT" variable in equation (1) is shown in the table and those are the discontinuities due to the age-dependent threshold. The discontinuities are shown on the basis of constant, linear, and quadratic polynomial terms in the time function and different window length to show the robustness of the results. The outcome variable is an employment dummy. And the left and right panels are based on the highest qualifications of individuals.

It suggests that the increase in minimum wage doesn't incur any changes in employment probability for individuals below GCSE who are the lowest levels of education among the sample. But qualifications below GCSE are very ambiguous since the qualifications are very diverse. There is a most robust positive effect for individuals with GCSE. Individuals with GCSEs have higher employment probability after becoming eligible for the higher minimum wage rate. The following results show that the significance mainly comes from individuals with higher numbers or better scores of GCSEs.

<Table 4.1 Here>

Table 4.2 presents the results of individuals with GCSEs as their highest qualification. The significant results concentrate on the individuals with higher qualifications. And the results are more significant on constant and linear terms. With higher polynomial terms, there is a risk of over-fitting. Individuals with five or more GCSEs have more employment probability after they become eligible for higher minimum wage rate, but there is no significant effect on individuals with less than five GCSEs. It suggests that increasing minimum wage will result in higher skilled workers getting into employment. There is a significant difference between the individuals with a higher number of GCSEs and a lower number of GCSEs. Since the level of education is very close for these two groups of people, their motivation for finding a job should be identical. It may imply that lower skilled-workers have less advantage for being employed after increasing the minimum wage, making them less attractive. To further testify the results, I also divided individuals regarding their grades of GCSE. Individuals with A-C GCSE still have an edge in gaining employment opportunities compared to individuals with D-G GCSEs. This evidence may imply that there is a crowding out effect from individuals with higher or better qualifications.

<Table 4.2 Here>

Figure 4.2 shows the predicted employment probability given different periods for individuals with higher numbers and grades of GCSE and it is calculated by

the mean of individual's employment probability given distances. The upper figure is the employment probability of individuals with five or more GCSEs. The lower figure is the employment probability of individuals with A-C GCSEs. Across the threshold, there is a clear jump of about 5% caused by the increasing minimum wage.

<Figure 4.2 Here>

Figure 4.3 shows the employment probability of individuals with less than five GCSEs and D-G GCSEs. The evidence of individuals with lower levels of education is not as strong as for higher levels of education. The effect might be a combined composition effect of the increasing minimum wage and crowding out effect.

<Figure 4.3 Here>

Table 4.3 gives estimates of the discontinuity in terms of number of GCSEs held by employed individuals and proportion of 5+ GCSEs. It is not an exact number of GCSE but a categorical variable associated with the number of GCSEs. Compared to the previous results, it shows the crowding out effect directly. Across the cut-off point, both results show a slight positive discontinuity, suggesting that individuals in employment on the right-hand-side of the cut-off holding more GCSEs compared to the left-hand-side. But the results are not significant compared to the employment probability results. It may be due to the magnitude of the crowding out effect.

<Table 4.3 Here>

4.4.2 Which job?

The different types of a job may also indicate the relative employment strength. It is crucial for young workers since the type of job has a significant influence on their return and stability, as well as the accumulation of human capital. Investigating the type of job will deepen the understanding of the employment probability across different groups. I investigate the probability of finding a full-time or a permanent job after becoming eligible for the higher minimum wage rate. Firstly, I examine the employment probability of finding a full-time job. But a full-time job can be a temporary as well and a permanent position could be a part-time job either. Moreover, I examine the employment probability of finding full-time permanent job.

Figure 4.4 is the proportion of individual in a full-time job given qualifications, like Figure 4.1. In Figure 4.4, individuals having qualifications less than GCSE and A-level don't have significant difference across the cut-off point. But for individuals with lower and higher numbers of GCSEs, there are clear trends for these two groups. Individuals with higher numbers of GCSEs have a higher probability of finding a full-time job after becoming eligible for higher minimum wage rate. Similar to the results of employment probability, the proportion of individuals with fewer numbers of GCSEs of finding a full-time job decreases after becoming eligible for higher minimum wage rate. Figure 4.5 shows the proportion of individual having a full-time permanent job. The pattern is quite similar to Figure 4.4.

<Figure 4.4 Here>

<Figure 4.5 Here>

From Table 4.4 to 4.6 I present the results of employment probabilities for different types of job given qualifications. There is more evidence for higher skilled workers finding a full-time permanent job compared with lower skilled workers. Together with Figure 4.4 and 4.5, I conclude that higher-skilled workers tend to find a full-time permanent job after the increase in minimum wage. However, there is no significant increase in the proportion of full-time permanent job for lower-skilled workers. Given the assumption I described above, both types of worker have similar motivation toward to a more full-time permanent job. The results of lower-skilled workers could be explained by a composition effect, combing with the effect of the increasing labour supply and crowding out effect.

<Table 4.4 Here>

<Table 4.5 Here>

<Table 4.6 Here>

4.4.3 Robustness and sensitivity check.

In order to test the sensitivity of the discontinuity, I run regressions based on different window lengths and polynomial terms. The more robust results are also significant after changing the window lengths and adding different polynomial terms.

To further check the sensitivity of the above results, I examine the effect on male and female separately. The results are shown in the appendix.

<Table 4.7 Here>

From Table 4.7, the discontinuity is more pronounced for males than for females. The pattern is also evident in the following figures of employment rate and predicted employment probability.

One of the purposes of age-dependent minimum wage is to allow the employers to discriminate workers by their age to minimize the cost. In another word, employers have the motivation to replace the older workers with the younger. In order to test the replacement effect, I run regressions on the basis of discontinuity of both -1 (d(-1))and -2 (d(-1)) months before the cut-off point where they become eligible for the higher minimum wage rate. The discontinuities are shown in Table 4.8. It is expected that the discontinuity will have a negative value if there exists displacement effect since employers will make workers redundant shortly before workers become eligible for the higher minimum wage rate. Here I examine the discontinuity for A-C GCSEs and D-G GCSEs followed by previous results. The results show that there is no significant discontinuity in two groups around the age threshold. It suggests that there are not strong replacement effects existing one month or two months before the threshold directly. A more indirect result suggests that crowding out effect starts one month before the cut-off point and employment probability doesn't have any discontinuity before threshold, because lower-skilled workers may be made redundant to some extent.

<Table 4.8 Here>

Moreover, to test the robustness, I also did the regressions on 21 and 23 years old to examine the discontinuity which should not exist. It increases more creditability to the story.

<Table 4.9 Here>

There were several policies coming out during the sample periods. Those may have subtle effects on the unemployment rate. Especially the unemployment rate in 2008 is still much lower compared to 2009 when it peakedand the higher level of unemployment rate holds constant until 2013.

4.4.4 Working hours.

Employees and employers will negotiate working hours based on each other's needs. It may also reflect the relation between each other. By the nature of RD, we can observe the discontinuity mainly caused by employees.

<Table 4.10 Here>

From the results of Table 4.10, we can see that an increase in minimum wage doesn't have a significant effect on lower-skilled workers. However, male workers with higher levels of skill tend to increase working hours and there is no significant discontinuity for females. It is also shown in Figure 4.6 which pictures the average working hours based on their qualification and distance to the threshold.

<Figure 4.6 Here>

More working hours of higher skilled workers may be caused by the higher proportion of full-time job among them. The interesting part is there is surprisingly no significant increase in working hours among lower-skilled workers considering that they have lower probabilities ending with a full-time job.

One possible explanation is that the increase in working hours of a part-time job has been offset by a lower full-time participation rate. Another possible explanation is that their working hours have already met both workers and employer's full capacity. Even though they tend to have higher wage after becoming eligible for the higher minimum wage, they can't increase working hours to increase income further. That implies that they may also want to find a full-time job with the higher expected return, but it is hard for them to find one due to higher competition after the increasing minimum wage.

One has to ask if minimum wage will effectively decrease income inequality? Even though their employment probability may not be affected significantly, the working condition and benefit may also be affected, especially in a bad time.

4.5 Conclusions.

This chapter enhances the literature by examining the effect of becoming eligible for adult minimum wage rate on employment probability for different groups and argues that the discontinuities are heterogeneous and the possible crowding out effect, caused by the increasing minimum wage. There is a strong evidence of heterogeneous effects by qualifications. The results suggest that there are no significant effects of an increasing minimum wage on the employment probability for individuals below GCSE and the significant discontinuity mainly comes from higher skilled workers. Individuals with higher numbers or grades of GCSE have a higher probability of being employed after the minimum wage increase. There are no significant effects for individuals with lower numbers or grades of GCSE. The number of GCSEs held by employed individuals and the proportion of 5+ GCSEs among employed workers increase across the threshold, suggesting that there is a crowding out effect. Moreover, the results also don't show that there exists significant replacement effect shortly before the age threshold.

Besides the general improvement in employment opportunity, the results also strongly suggest that higher-skilled workers tend to find a full-time permanent job after becoming eligible for higher minimum wage rate. The evidence of higher employment probability and higher satisfying job accession probability may imply that there is a crowding out effect coming from higher skilled workers. On the other hand, there is no evidence in terms of lower-skilled workers being made redundant.

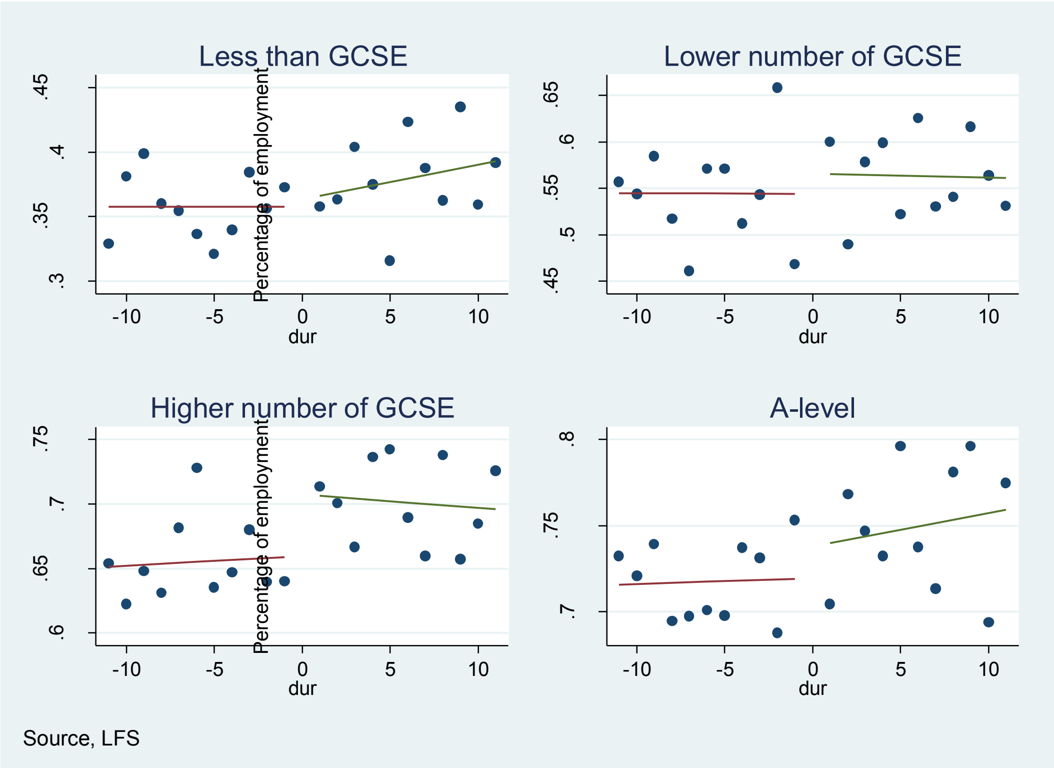
However, the limited information of the dataset does not allow me to examine the replacement effect in more details or labour flows in general. With data in a booming period, one might expect more variation between the local unemployment rate and the employment probability.

The implication of this chapter is that we should not neglect the possible adverse effects of the increase in minimum wage, especially in a recessionary period. Due to a tighter labour market, the discontinuity caused by increasing minimum wage will not only increase labour supply but also result in less chance for disadvantaged workers in labour market. Moreover, even though the negative effect of minimum wage is limited on average, the negative effect in subgroups may be still non-negligible. The results imply that exogenous increase in competition may have both immediate and long-term negative effects on disadvantaged workers. The minimum wage policy should be more flexible in a tighter labour market.

More work is needed in the future. The RD design could be more flexible and convincing if day of birth is available to construct the distances. Moreover, the stock of employment could be ambiguous. It might be clearer to examine crowding out effect by examining the flow of labour, but that requires a more comprehensive dataset.

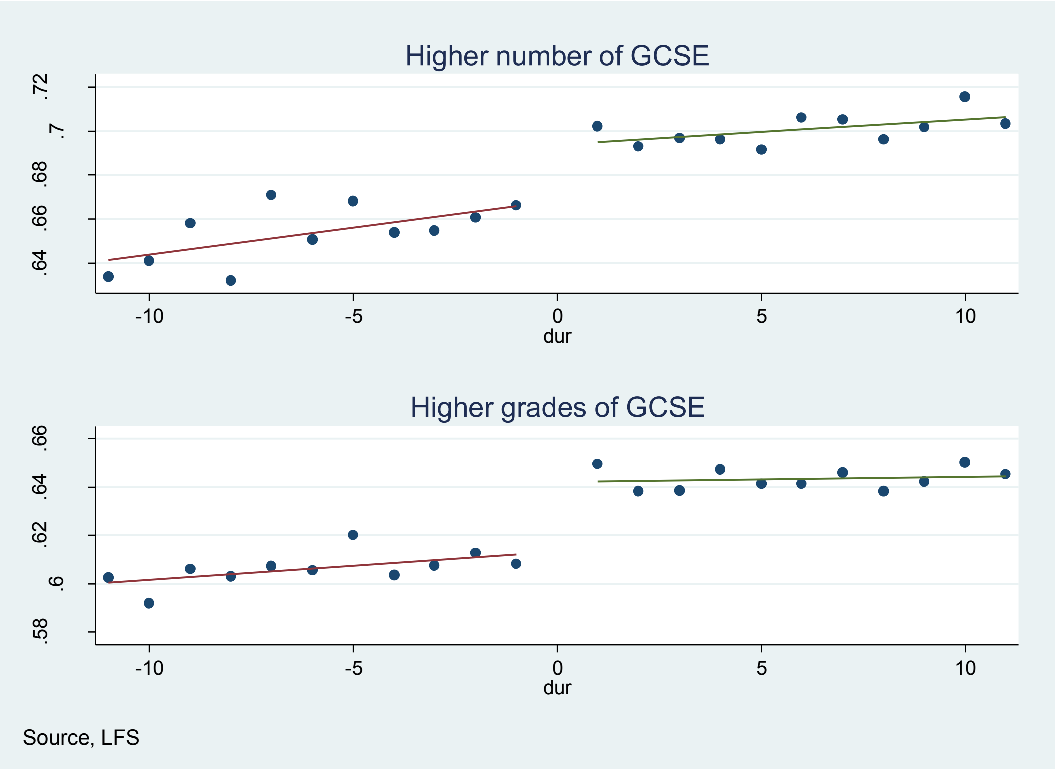
Figures and Tables:

Figure 4.1 Probability of being employed by qualifications



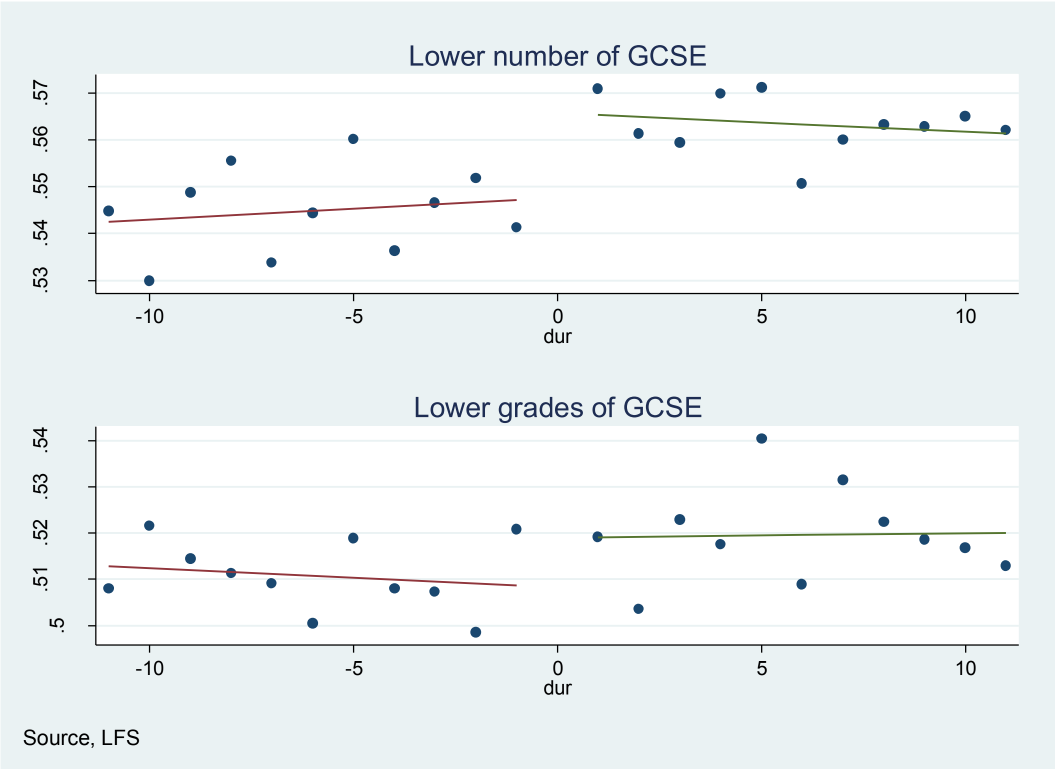
Notes: The variable "less than GCSE" includes individuals whose highest qualification is lower than GCSE according to variable 'HIQUAL' in QLFS. Lower number of GCSE includes individuals whose highest qualification is GCSE and hold less than 5 GCSEs. Higher number of GCSE includes individuals whose highest qualification is GCSE and hold five or more GCSEs. Distances include -11 to 11 from the month of survey to the month of birth.

Figure 4.2 Estimated probability of employment for A-C or 5+ of GCSE



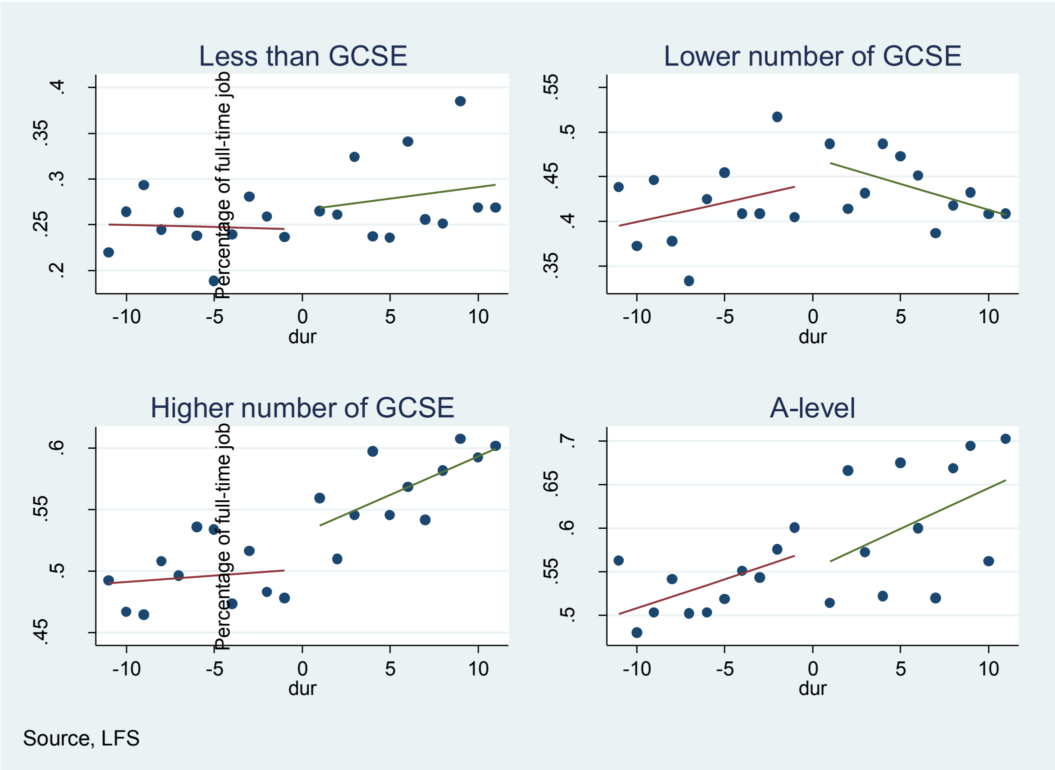
Notes: The Figure plots the estimated average probability of employment rate which is calculated by averaging individual's estimated employment probability on the basis of distances.

Figure 4.3 Estimated probability of employment for D-G or 5- GCSE



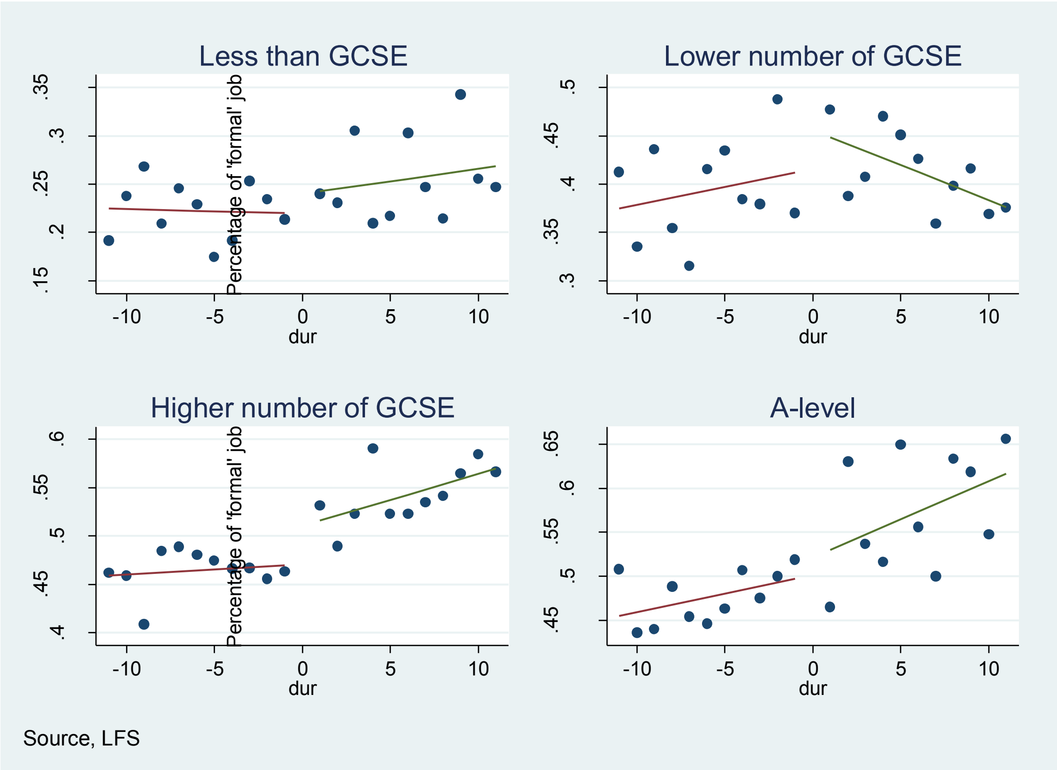
Notes: The Figure plots the estimated average probability of employment rate that is calculated by averaging individual's estimated employment probability on the basis of distances.

Figure 4.4 Employment rate for individuals with a full-time job



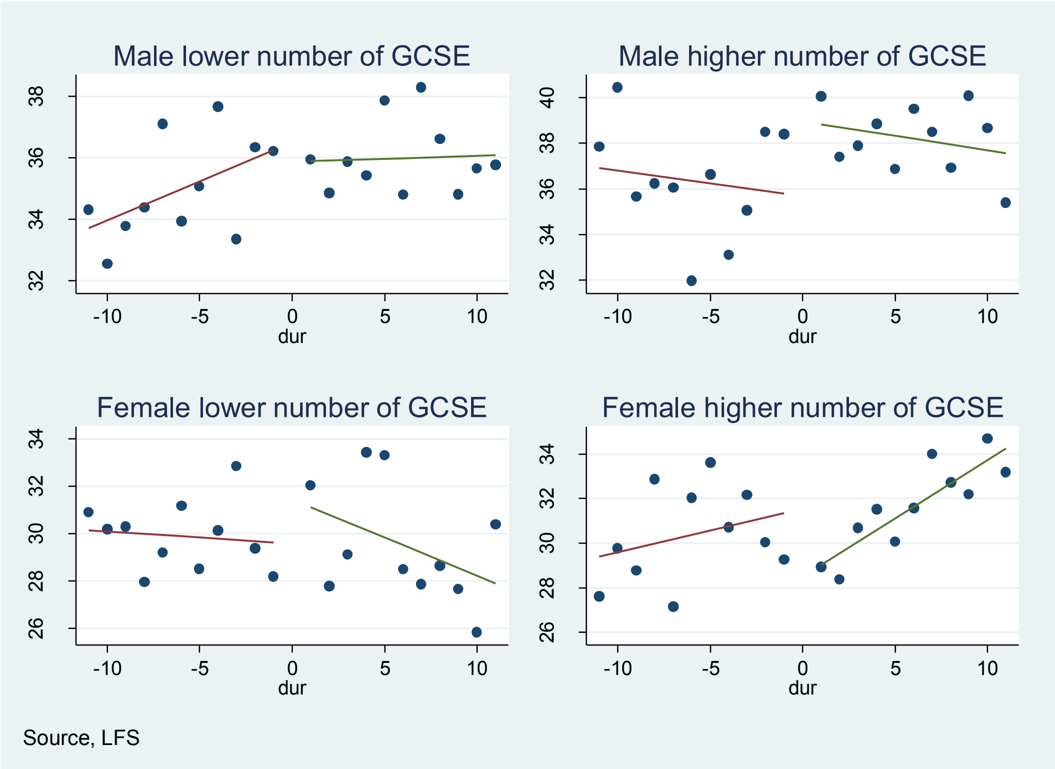
Notes: The Figure plots the employment rate of full-time job. Sample includes inactive, unemployed, Part-time job, and Full-time job.

Figure 4.5 Employment rate for individuals with a full-time permanent job



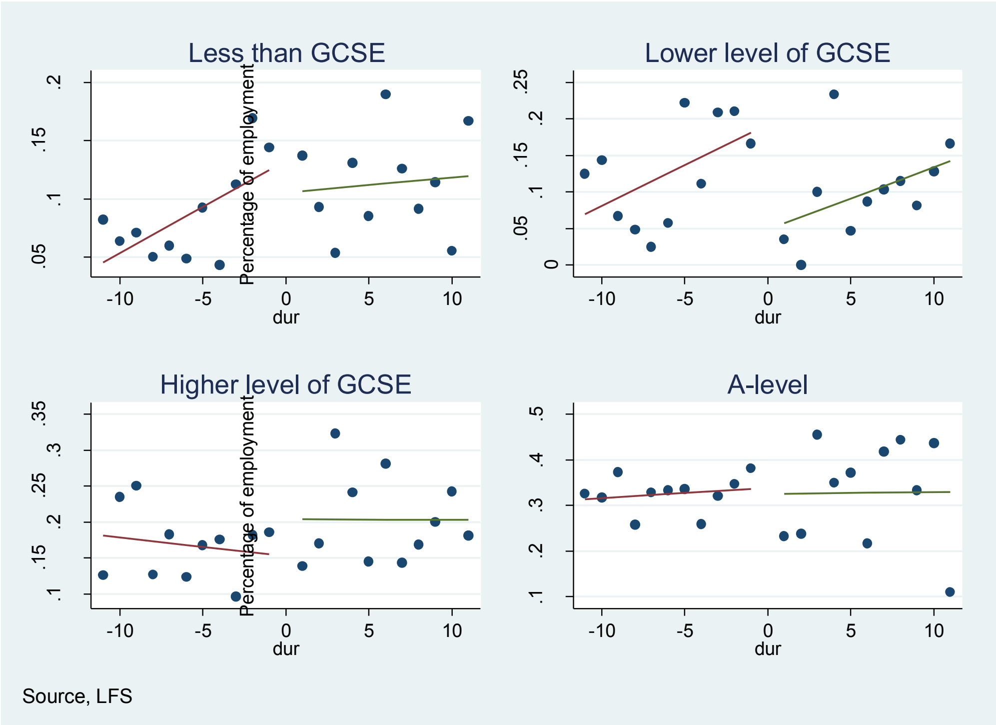
Notes: The Figure plots the employment rate of both full-time and permanent job. All samples are included.

Figure 4.6 Working hours of 5- or 5+ workers



Notes: The Figure plots the working hours in different types of qualification. Only GCSE workers are considered.

Figure 4.7 Proportion of higher levels occupation



Notes: Distances are calculated by the subtraction between month of job started and month of birth at the 21nd birthday. Students are not included. All observations are employed.

Table 4.1 Employment probability of GCSE and below GCSE

	Employ (less than GCSE)			Employ (GCSE)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.069*	0.062	0.001	0.080**	0.084**	0.448
	[0.037]	[0.050]	[0.088]	[0.030]	[0.040]	[0.071]
Linear	0.025	-0.074	-0.065	0.107*	0.088	0.686**
	[0.080]	[0.115]	[0.283]	[0.064]	[0.092]	[0.228]
Quadratic	-0.054	0.017		0.057	0.181	
	[0.136]	[0.222]		[0.109]	[0.180]	
Observation	4842	2618	638	7297	3968	1332

Notes: The dependent variable is whether individual is in employment. The category of less than GCSE includes individuals with qualification lower than GCSE according to variable of "HIQUAL" in QLFS and GCSE includes all individuals whose highest qualification is GCSE. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.2 Employment probability of subgroups

	Employ (5- GCSEs)			Employ (5+ GCSEs)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.045	0.037	-0.067	0.119**	0.150**	0.230**
	[0.038]	[0.052]	[0.095]	[0.048]	[0.066]	[0.111]
Linear	0.050	-0.011	1.10***	0.193*	0.225*	0.277
	[0.083]	[0.121]	[0.309]	[0.104]	[0.148]	[0.362]
Quadratic	-0.060	0.267		0.271	0.087	
	[0.142]	[0.235]		[0.176]	[0.286]	
Observation	4365	2357	759	2932	1611	573

	Employ (GCSEs D-G)			Employ (GCSEs A-C)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	-0.001	-0.065	-0.244	0.098***	0.119**	0.117
	[0.068]	[0.094]	[0.165]	[0.033]	[0.045]	[0.079]
Linear	-0.118	-0.257	0.607	0.171**	0.193*	0.073**
	[0.145]	[0.212]	[0.547]	[0.072]	[0.104]	[0.257]
Quadratic	-0.340	0.135		0.181	0.206	
	[0.247]	[0.410]		[0.123]	[0.202]	
Observation	1400	755	119	5897	3213	1077

Notes: The dependent variable is whether individual is in employment. The category of 5- GCSE includes individuals with less than 5 GCSE and 5+ GCSE includes individuals with five or more GCSE. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.3 Crowding out effect

	Number of GCSE			Proportion of 5+ GCSEs		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.052**	0.064**	0.067	0.048**	0.048*	0.026
	[0.023]	[0.032]	[0.055]	[0.019]	[0.026]	[0.046]
Linear	0.077	0.095	-0.025	0.047	0.023	-0.080
	[0.051]	[0.073]	[0.177]	[0.042]	[0.061]	[0.148]
Quadratic	0.091	-0.052		-0.009	-0.086	
	[0.086]	[0.142]		[0.072]	[0.118]	
Observation	16631	9060	3026	16631	9060	3026

Notes: The dependent variable are the full time of numbers of GCSE and if individuals hold five or more GCSE. Only employed observations are included and the observations are restricted with less than 20 years continuous education in order to get rid of the effects of new graduates who have much higher years of education. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.4 Job type of individuals with less than GCSE and GCSE

	Full (less than GCSE)			Permanent-Full (less than GCSE)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.120**	0.124**	0.064	0.123**	0.122**	0.053
	[0.039]	[0.053]	[0.095]	[0.040]	[0.058]	[0.097]
Linear	0.073	-0.016	0.240	0.075	0.018	0.277
	[0.084]	[0.123]	[0.304]	[0.086]	[0.125]	[0.312]
Quadratic	0.038	0.254		0.068	0.155	
	[0.144]	[0.234]		[0.147]	[0.238]	
Observation	4842	2618	864	4842	2618	864

	Full (GCSE)			Permanent-Full (GCSE)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.106***	0.094**	0.042	0.111***	0.111**	0.060
	[0.030]	[0.040]	[0.071]	[0.030]	[0.040]	[0.071]
Linear	0.088	0.064	0.428*	0.117*	0.098	0.477
	[0.064]	[0.092]	[0.229]	[0.091]	[0.092]	[0.228]
Quadratic	0.022	0.095		0.055	0.093	
	[0.108]	[0.178]		[0.109]	[0.179]	
Observation	7292	3968	1332	7292	3968	1332

Notes: The dependent variable is whether individual is in full-time job or both full-time and permanent job. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.5 Job type of individuals with lower numbers or grades of GCSE

	Full (D-G GCSE)			Permanent-Full (D-G GCSE)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	-0.003	-0.073	-0.165	0.14	-0.052	-0.085
	[0.069]	[0.096]	[0.167]	[0.070]	[0.096]	[0.168]
Linear	-0.109	-0.233	-0.017	-0.058	-0.056	0.106
	[0.149]	[0.218]	[0.560]	[0.150]	[0.219]	[0.563]
Quadratic	-0.257	0.004		-0.158	0.071	
	[0.252]	[0.419]		[0.252]	[0.425]	
Observation	1400	755	255	1400	755	255

	Full (5- GCSEs)			Permanent-Full (5- GCSEs)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.044	0.053	-0.034	0.044	0.064	0.003
	[0.039]	[0.053]	[0.096]	[0.039]	[0.053]	[0.095]
Linear	0.060	-0.026	0.688**	0.091	0.041	0.797**
	[0.084]	[0.123]	[0.311]	[0.084]	[0.123]	[0.311]
Quadratic	-0.079	0.099		-0.008	0.155	
	[0.144]	[0.238]		[0.144]	[0.238]	
Observation	4365	2357	388	4365	2357	388

Notes: The dependent variable is whether individual is in full-time job or both full-time and permanent job. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.6 Job type of individuals with higher numbers or grades of GCSE

	Full (A-C GCSE)			Permanent-Full (A-C GCSE)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.130***	0.130**	0.092	0.132***	0.144***	0.091
	[0.033]	[0.045]	[0.079]	[0.033]	[0.066]	[0.079]
Linear	0.137**	0.156	0.552**	0.161**	0.145	0.572**
	[0.071]	[0.103]	[0.256]	[0.071]	[0.103]	[0.256]
Quadratic	0.128	0.143		0.116	0.121	
	[0.121]	[0.199]		[0.121]	[0.199]	
Observation	5897	3213	524	5897	3213	524

	Full (5+ GCSEs)			Permanent-Full (5+ GCSEs)		
	22 months	12 months	4 months	22 months	12 months	4 months
Constant	0.179***	0.145**	0.177	0.190***	0.169**	0.166
	[0.047]	[0.064]	[0.110]	[0.047]	[0.064]	[0.109]
Linear	0.117	0.198	0.176	0.142	0.176	0.098
	[0.100]	[0.143]	[0.254]	[0.101]	[0.220]	[0.352]
Quadratic	0.194	0.115		0.168	0.033	
	[0.169]	[0.275]		[0.170]	[0.277]	
Observation	2932	1611	573	2932	1611	573

Notes: The dependent variable is whether individual is in full-time job or both full-time and permanent job. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.7 Employment probability for male and female

	Employ (less than GCSE)				Employ (GCSE)			
	Male		Female		Male		Female	
	22 months	12 months	22 months	12 months	22 months	12 months	22 months	12 months
Constant	0.078*	0.111	0.042	0.014	0.099**	0.104*	0.055	0.046
	[0.050]	[0.069]	[0.055]	[0.075]	[0.041]	[0.056]	[0.043]	[0.059]
Linear	0.067	-0.114	-0.017	-0.054	0.116	0.064	0.075	0.088
	[0.109]	[0.157]	[0.120]	[0.174]	[0.088]	[0.128]	[0.095]	[0.137]
Quadratic	-0.055	0.070	-0.078	-0.100	-0.001	0.103	0.094	0.209
	[0.185]	[0.302]	[0.205]	[0.333]	[0.150]	[0.248]	[0.163]	[0.269]
Observations	2569	1402	2273	1216	3944	2170	3353	1798

	Employ (5- GCSE)				Employ (5+ GCSE)			
	Male		Female		Male		Female	
	22 months	12 months	22 months	12 months	22 months	12 months	22 months	12 months
Constant	0.029	-0.015	0.058	0.080	0.238***	0.324***	0.012	-0.009
	[0.051]	[0.071]	[0.058]	[0.080]	[0.071]	[0.096]	[0.068]	[0.093]
Linear	-0.014	-0.037	0.106	0.008	0.334**	0.186	0.059	0.317
	[0.110]	[0.164]	[0.128]	[0.188]	[0.151]	[0.210]	[0.146]	[0.213]

Quadratic	-0.133	-0.009	0.015	0.578	0.188	0.193	0.338	-0.042
	[0.191]	[0.318]	[0.221]	[0.367]	[0.250]	[0.403]	[0.252]	[0.423]
Observations	2497	1364	1868	993	1447	806	1485	805

	Employ (D-G GCSE)				Employ (A-C GCSE)			
	Male		Female		Male		Female	
	22 months	12 months	22 months	12 months	22 months	12 months	22 months	12 months
Constant	-0.029	-0.180	0.082	0.149	0.144***	0.187***	0.047	0.028
	[0.089]	[0.122]	[0.110]	[0.154]	[0.047]	[0.064]	[0.048]	[0.065]
Linear	-0.272	-0.287	0.165	-0.151	0.238	0.184	0.080	0.173
	[0.188]	[0.271]	[0.241]	[0.361]	[0.101]	[0.147]	[0.104]	[0.150]
Quadratic	-0.407	0.051	-0.247	0.305	0.134	0.128	0.208	0.223
	[0.320]	[0.526]	[0.424]	[0.714]	[0.172]	[0.285]	[0.178]	[0.295]
Observations	840	457	560	298	3104	1713	2093	1500

Notes: The dependent variable is whether individual is in employment. The category of 5- GCSE includes individuals with less than 5 GCSE and 5+ GCSE includes individuals with five or more GCSE. A-C GCSE and D-G GCSE stand for individual's highest qualification is A-C or D-G. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.8 Displacement effect

VARIABLES	D-G GCSE d(-1)	A-C GCSE d(-1)	5- GCSE d(-2)	5+ GCSE d(-2)
discontituity	-0.167	0.004	0.156	0.083
	(0.147)	(0.073)	(0.158)	(0.078)
duration	0.009	0.009	-0.0215	0.001
	(0.017)	(0.009)	(0.020)	(0.010)
Duration discontinuity	0.008	-0.009	0.028	-0.003
	(0.022)	(0.011)	(0.025)	(0.012)
Constant	1.965***	1.607***	1.682***	1.539***
	(0.481)	(0.225)	(0.490)	(0.229)
Observations	1,400	5,897	1,400	5,897

Notes: Probit model. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. The d(-1) and d(-2) denote the one month and two months before the age threshold.

Table 4.9 19-20 and 22-23 discontinuity

	20 years old				23 years old			
	Employ (A-C GCSE)		Employ (D-G GCSE)		Employ (A-C GCSE)		Employ (D-G GCSE)	
	22 months	12 months	22 months	12 months	22 months	12 months	22 months	12 months
Constant	0.025	-0.002	0.068*	0.040	0.049	0.071	0.045	0.013
	[0.032]	[0.044]	[0.037]	[0.050]	[0.035]	[0.048]	[0.069]	[0.094]
Linear	-0.050	-0.052	0.008	0.048	0.075	0.055	0.043	0.234
	[0.070]	[0.101]	[0.081]	[0.116]	[0.075]	[0.108]	[0.147]	[0.215]
Quadratic	-0.059	-0.035	0.070	0.132	0.087	0.052	0.262	0.181
	[0.119]	[0.196]	[0.138]	[0.225]	[0.127]	[0.206]	[0.252]	[0.409]
Observations	6342	3401	4571	2502	5495	2980	1337	730

Notes: The dependent variable is whether individual is in employment. The category of 5- GCSE includes individuals with less than 5 GCSE and 5+ GCSE includes individuals with five or more GCSE. A-C GCSE and D-G GCSE stand for individual's highest qualification is A-C or D-G. The results are based on different polynomial terms and window length. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Table 4.10 Working hours among male employed workers

	Working hours							
	Male				Female			
	Employ (5+ GCSE)		Employ (5- GCSE)		Employ (5+ GCSE)		Employ (5- GCSE)	
	22 months	12 months	22 months	12 months	22 months	12 months	22 months	12 months
Constant	1.551**	1.753**	0.096	0.483	0.832	-1.451	-0.676	-0.366
	[0.575]	[0.803]	[0.608]	[0.781]	[0.762]	[1.029]	[0.817]	[1.169]
Linear	2.059*	-0.190	-0.754	-1.668	-2.660	0.576	0.043	0.848
	[1.127]	[1.469]	[1.272]	[1.773]	[1.672]	[2.441]	[1.992]	[2.901]
Quadratic	-1.084	0.494	-1.161	-1.510	0.129	-2.901	1.912	1.937
	[1.656]	[2.787]	[2.146]	[3.396]	[2.918]	[4.666]	[3.641]	[6.252]
Observations	1027	573	1553	866	960	531	864	461

Notes: The dependent variable is working hours in every period. Control variables are local unemployment, dummy for disable, month of birth, and years. * significant at 10%; **significant at 5%; ***significant at 1%;

Chapter 5

Quantitative effects of higher education expansion on the returns to education: Evidence from the UK

5.1 Introduction and literature review.

5.1.1 Introduction.

After the 'Education Reform Act 1988', increasingly graduates left full-time education with a degree. The proportion of university graduates increased from 15% for men and 13% for women to 30% and 35% respectively (Walker and Zhu, 2008). Not only the number of graduates increased due to the reform but also the average level of education. However, the increased average education is not solely due to the reform. It was increasing rapidly in the recent decades. The people who hold the A-level qualification increased significantly and on the other hand the number of individuals who hold GCSEs as the highest qualification fell. Although due to the endeavour of the British government to vocational education such as National Vocational Qualification (NVQ), the numbers of vocational student increased slightly during that periods.

The reform allowed universities to relax the recruiting requirements for the students. The newly recruited university graduates might be less capable compared to the previous graduates as a result of the relaxed requirements, leading to ambiguous results if the compositional changes are not controlled. The literature minimizes ability bias by narrowing down the scope in which only includes individuals with at least one or two A-levels (Blundell et al, 2000; Walker and Zhu, 2008). Walker and Zhu (2008) examine how the education expansion affects university wage premium. Interestingly, they don't find strong negative effects on the returns of new graduates and even marginal positive effect for women.[75][76] Subsequently, Devereux and Fan (2011) argued that because of the increasing number of individuals with A-levels, the results may also be biased. They apply Two-Stage Lease Squared (TSLS) to examine the wage premium of a degree using Quarterly Labour Force Survey (QLFS) and argue that the increasing education has a 6% wage premium for both men and women. However, there are returning students who pursue a degree as a result of the reform. In this study, the returning students represent the student who the age of obtaining highest qualification is larger than the age of leaving full-time education. Their TSLS results may also lead to a misunderstanding due to this heterogeneity. If the increasing numbers of returning student become university graduates due to the reform, their results would be biased due to this additional channel that they haven't picked up. They did not check how relaxed entry requirements affected the opportunity of attending universities. Walker and Zhu (2011) argue that the returns vary a lot by subjects and the rise of tuition fees

[75] They point out that the results might be biased by the personal innate ability. They also examine the heterogeneity based on a quantile regression.

[76] An important point for DID analysis is to select the comparable control group. The results may be completely different with different control groups.

had the relatively lower effects on the overall return for students in the UK. Lindley and Mcintosh (2015) firstly examined income inequality among graduates. They applied a variance analysis and argued that the widening variance of personal abilities is because of the differential accepting rules among universities. They argue that the large income inequalities among graduates mostly come from differences within a subject rather than between subjects. They also show the relation between the relaxed university entry requirements and the large wage variation. The literature implies there exists strong heterogeneities among the returns of university graduates.

Due to the fact that the reform may lead to more students with distinctive educational background entering into universities, the results might be very heterogeneous. The relaxed requirements in universities may not only open doors for the fresh students but also individuals who have a similar level of education and several years of working experiences, namely the returning students. It is commonly believed that individuals with a lower level of education tend to have a lower innate ability. However, Birch and Miller (2007) find that students who like to defer universities are found with higher schooling marks compared to those who start university right away from high schools. Their results and the fact that abilities tend to be multiple dimensions may imply this question might be more complex than we expected. Due to the limit of the data, previous studies fail to examine this heterogeneity. I provide evidence of the heterogeneity of the effect induced by the education expansion.

The current literature argues that the education expansion doesn't incur largely negative effects on the returns to the newly recruited graduates. Surprisingly the effects of the education reform on the returning students have not been examined. Individuals with working experiences tend to be older and more mature compared to other younger students. They may arrange their learning plan and also the career plan after study better. More importantly, they may choose the most suitable learning plan combining with their existing human capital and future career plan. On the other hand, they tend to have more years of working experiences, but they also tend to have fewer years of fundamental education.[77] To my knowledge, there are very few papers which focus on the determinants of the decision of re-education, the effect of re-education or the effect of education expansion on returning students. Not only those A-levels graduates will comply with this reform, but also some returning students who are willing to obtain more education will also comply the education expansion. One of the critics of vocational education is its weak link to higher education. If those newly recruited returning graduates are successful after obtaining a degree compared to their previous colleagues, then the fact may shed light on how to conduct education reform in the future. This group of people could be very heterogeneous and therefore need a very informative dataset. But analysing the effect of obtaining more education on this group is of particular public policy

[77] They tend to have few years full-time education.

interest.

5.1.2 The contribution.

This chapter mainly focuses on examining the quantitative effect of the increasing university graduates on returns. Moreover, I illustrate another implicit achievement of the education expansion. That is to provide opportunities not only for the fresh students but also for the returning students who may benefit more from becoming university graduates out of clear purposes. I contribute to the literature in two ways.

First, the 2SLS and statistical results show that the returns are rather heterogeneous. The individual with a lower probability of attending a university may benefit more compared to the newly recruited graduates. However, they may come from specific reasons, namely the unobservable factors. From the results, it is clear that the return to university is negatively correlated with the propensity score of the attending universities. From Figure 5.7, the trends are very clear that individuals with lower probabilities of attending university have higher working experiences and lower education, holding lower numbers of A-levels and GCSE. To sum up, there are three types of new university graduates. They are returning students, A-level fresh students and higher level of fresh vocational education students respectively.[78] Due to the heterogeneous background among the new

---

[78] From the results, most newly recruited university graduates are with less than 24

university entrants and complex measurement errors such that individuals may report as a university student since they may attend a Polytechnic before the reform.[79] The instruments to estimate the propensity scores are birth cohorts and the instruments are considered to be exogenous and meet the exclusion restriction (Devereux and Fan, 2011). Moreover, the subjects are included as control variables in the reduced form to control for the underlying heterogeneity. Due to the fact that different subjects have different requirements, it may lead to the bias into the results if it is not controlled for (Lindley and Mcintosh, 2015; Walker and Zhu, 2011).

Second, I apply the simple Difference-in-Difference (DID) methodology to estimate the effect of an increase in the supply of university graduates on the returns directly and Matching Difference-in-Difference (MDID) to correct for the bias. The propensity score is the probability of attending university.[80] I apply two different methods to perform the matching strategy. One is the popular propensity score matching. Another is called Coarsened Exact Matching which it has become popular in recent years in order to check the robustness. Compared to the previous studies, I don't focus on individuals who have at least one A-level due to the fact that increasingly returning students born after 1976 obtain a degree. It will bring unnecessary bias into the results if we drop individuals who

continuous years of education, see Figure 5.1.

[79] Further and Higher Education Act 1992.

[80] The propensity score is estimated by "Probit" model, based on year of birth, numbers of A-level, numbers of GCSE, industry, sex, years of survey, quarter of survey education, squared of education, experiences, squared of experiences, tenure, marriage, job training, disable, and London dummy.

don’t have A-levels. The matching strategy will balance the characteristics between the treatment group and control group before and after the reform. It corrects the ability bias and estimates the quantitative effect of the increasing graduates on the returns. As I mentioned above, due to the limitations of the data, I can’t capture the change in the scores at A-level. Since the data doesn’t include information regarding the exact numbers of A-level and the grades of A-levels, the changes in the composition of A-levels are ambiguous to identify new graduates who enter into university without work experiences.

In order to best account for the heterogeneous effect, I divided the sample into three different periods, before-expansion, during-expansion, and post-expansion based on their birth cohorts (Devereux and Fan, 2011). Individuals who are born between 1970 and 1975 belong to “pre-expansion”. Individuals who are born between 1970 and 1975 belong to “during-expansion” Individuals who are born after 1975 belong to “post-expansion”. After correcting the ability bias, the results show that there is a strong negative effect for the fresh graduates in the post-expansion period. There are around 46% university graduates who obtain a degree after finishing the continuous education. From the DID results, both fresh students and returning students don’t have significantly negative effects in both periods. The MDID results correct the innate ability bias and shows that there exists significant negative effect on returns for the fresh students. Here, the MDID results for the returning students may correct the changes in the characteristics in which is captured by age, years of education and numbers of A-levels. The MDID results for the fresh students are different from the DID results in post-expansion period because there are increasing individuals who have more years of education.[81]

5.1.3 Source of Variation of matching-DID.

The problem with the current literature is that few literature shed light on how the characteristics changed among newly recruited university graduates. In another word, the bias comes from the change in the innate personal ability, leading to a non-identical treatment group as a result of the reform. The simple DID results combine qualitative effect with quantitative effect.

The numbers of the A-level can't fully capture the change in the admission rules for the A-level graduates. But the compositional change may capture the change in relative probabilities of attending universities for the returning students and higher levels of vocational education students. The decrease in numbers of A-levels may come from vocational students who have similar years of education with the fresh students. The compositional changes would be captured in the numbers of A-level, years of continuous education or years of working

[81] From Figure 5.7, we can see that the years of education increase continuously along with the propensity score. Individuals with the highest propensity score might be students with higher years of education than A-level graduates, but obtain the degree due to the reform.

experience. Those variations may also be captured in MDID results. That may explain why there is a difference between DID and MDID results for fresh students.

5.1.4 Potential biases and future work.

The first one is the incapability of capturing the compositional change in terms of personal abilities, such as of score and number of A-levels. Due to the fact that the parental background has impacts on higher education attainment, parental background might be helpful to control the compositional change. By taking advantage of the Understanding Society, the results are very similar to the QLFS. Dearden (1999) argue that the OLS result can be 30% upward biased due to omitted variables of personal ability or background. For the returning students, the newly recruited university graduates might be different from those who don't participate in universities.

Second, universities could be very heterogeneous in terms of the types of qualification, the quality of the programs, the reputation of the institutions, and so on. It is expected that some returning students will take a part-time degree that can't be captured in my data and also the quality of institutions (Walker and Zhu, 2011). The third is that there is a lack of information regarding the educational background of newly recruited graduates. In the current data, only their continuous years of education, numbers of A-level and numbers of GCSE are available to me. But I don't have the information in terms of their highest education before obtaining a degree. This could lead to ambiguous results.

Regardless of the above potential biases, there are several interesting extensions can be done. The returning students are both politically and personally interesting to explore. In this study, I present the heterogeneous returns compared to the fresh students. However, due to the limit of data, I couldn't examine the effects of characteristics of determining to pursue a degree, which presumptively it will lead a selection bias when examines the returns among returning students. It needs a very informative dataset which includes a comprehensive background. After exploring why they decide to pursue a degree and what is the outcome of that, more policies could be made in order to differentiate the market and to provide more efficient courses for the returning students.

5.2 Data.

5.2.1 Data description.

The main data is drawn from the 2002 to 2013 Quarter Labour Force Survey. The birth cohorts are from 1965-1979 and age range of the sample is from 33 to 43 years-old since observations are only matched in this age band, shown in Figure

5.1.[82] This feature has an advantage that it can allow us to examine the full potential returns when they are in the middle of their career (Blundell et al, 2000). Another part of the data is drawn from Understanding Society wave A to wave E. The age range is from 40 to 45 years-old and the birth cohorts are from 1965-1975. The parental occupation is available in Understanding Society. It could be used to capture the background of the newly recruited university graduates.

<Figure 5.1 Here>

The numbers of individuals who hold A-levels are shown in Figure 5.2. The categorical variable in QLFS indicates if an individual holds one, more than two or none A-levels.[83] So here the y-axis shows the ambiguous numbers of A-level which individual holds given different birth cohorts. It clearly shows that the total number of individuals holding A-level has largely increased after 1970 birth cohort. It suggests that the higher education expansion has pushed the students to get more A-levels or the grade inflation.

<Figure 5.2 Here>

Figure 5.3 shows that the number of A-levels among graduates given birth cohorts. It shows that the proportion of A-levels among the graduates gradually decreased over time, indicating that there is a compositional change among the

---

[82] This study particularly focuses on the effects in England and Wales.
[83] "NUMAL" in QLFS refers to how many A-levels does an individual holds.

university graduates.

<Figure 5.3 Here>

The personal innate ability might be a serious bias on the basis of the simple DID. Previous literature neglects the compositional change among university graduates. It turns out to be a very important change along with the higher education expansion.[84] More returning students become university graduates, especially born after 1975. In the QLFS sample, the proportion of the returning student among university graduates is higher than the fresh students after 1978. That may potentially explain why there are robust heterogeneous returns among the graduates and why the returns in 1976-1979 are significantly lower compared to the period between 1970 and 1975. Figure 5.4 has shown this trend. After 1970, the numbers of university graduate of fresh students increased. However, the returning students tend to be constant during that period and it increased significantly after 1975.

<Figure 5.4 Here>

Figure 5.5 shows the individual's age when they completed the full-time education. There are around 41% university graduates who belong to the returning students. Among them, there are sufficient amount of the returning students with at least 20 years of continuous education before becoming

[84] Back to 1970s, students who go to the universities may have more homogeneous background because of the O-level and fewer universities.

university graduates. Most of the returning students have around 16 years, 18 years, and 21 years full-time education. Those may refer to A-level, GCSEs level, and HNC/HND level respectively.

<Figure 5.5 Here>

5.2.2 Measurement error.

First, it is a self-reported data. People may report their "university" status given the current classification, but their self-reported full-time years of education fail to include the years spending in the university. Second, it is a part-time degree. It takes more time compared to a full-time degree in which can't be controlled in this study. Moreover, I differentiate the returning students and the fresh students based on their self-reported continuous years of education and age when completed the qualification. The measurement error is inevitable. The robustness checks are necessary.

5.3 Matching Difference-in-Difference.

Due to the fact that universities have relaxed the entry requirements, the university graduates matriculated before 1988 may have different personal innate ability compared to the graduates matriculated after 1988. The simple

DID may be biased if those compositional change may enter the group-specific effects and those effects fail to control for. MDID is a promising method to tackle with this problem, although it also depends on the data which it includes variables describing the compositional change before they take the treatment (degree), such as scores or numbers of A-level. In my dataset, it only includes the number of A-levels, so intuitively it can only capture the compositional change among the returning students since they tend to hold fewer numbers of A-levels or even none. This method was originally developed by Heckman et al (1997, 1998).

There are mainly three types of biases. One is the selection on the un-observables. Another one is the failure of a common support condition and the last one is a failure to weight treatment and comparison group comparably for which they argue that it is unlikely happened in matching strategy (Heckman et al 1997). The first and the second bias will be corrected by the matching method. However, the common support may bring additional bias if treatment effect is heterogeneous among the treated group (Blundell et al, 2005).

D is the treatment. The university graduates are treated individuals, denoted as $D_i = 1$. Others remain in the control group, denoted as $D_i = 0$. Y_{it} denotes the outcome of individual i in time t, before the reform. $Y_{it'}$ denotes the outcome of individual i in time t', after the reform. $X_1 and\, U_1$ are the observables and un-observables for the treated group. $X_0 and\, U_0$ are the observables and un-observables for the control group. ATE is the average treatment effect.

$$ATE = E(Y_1|X_1, U_1) - E(Y_0|X_0, U_0) \quad (1)$$

$$B = E(U_1) - E(U_0)$$

And the bias will become zero when treatment assignment is independent conditional on X.

$$Y_1, Y_0 \perp D \mid X$$

That means $E(Y_1|X_1, D = 0) = E(Y_0|X_0, D = 1)$

Given the assumption of "Strong Ignorability" proposed by Rosenbaum and Robin (1985), 0<P(D=1|X)<1. Together with the former two equations, that implies the following,

$$(Y_0, Y_1) \perp D \mid P(X)$$

$$E(Y_1|P(X), D = 0) = E(Y_0|P(X), D = 1) \quad (2)$$

Essentially X can be decomposed into (T, Z). T is variables retained in the reduced form and Z is the exogenous variables in the first stage.

$$Y_1 = f_1(T) + U_1, \qquad Y_0 = f_0(T) + U_0$$

Due to the fact that Z is completely exogenous, it leads to $U_0 \perp D \mid Z$. With same spirit with "Strong Ignorability", it leads to $Y_1, Y_0 \perp D \mid P(Z)$. Then,

$$E(U_0|P(Z), D = 0) = E(U_0|P(Z), D = 1) \quad (3)$$

Matching is still a "selection-on-observables" method, under the language of Heckman and Robb (1985). Common support problem can be eliminated if matching is performed over common support. The Conditional Independence Assumption (CIA) doesn't hold when the unobservables affect the outcome even under the control of propensity score. MDID relaxes the CIA from single observation to pair-wise. In this setting, we only need the CIA holds in the first difference equation.

$$E(Y_{0t} - Y_{ot'}|X_0, D = 0) = E(Y_{0t} - Y_{ot'}|X_0, D = 1) \tag{4}$$

Under index sufficiency the equation becomes

$$E(Y_{0t} - Y_{ot'}|P(Z), D = 0) - E(Y_{0t} - Y_{ot'}|P(Z), D = 1) \tag{5}$$

The results normally vary with different matching scheme. Different weights have been proposed.

$$ATT = \frac{1}{N} \sum_{i \in I_1} [Q_{1i} - \sum_{j \in I_0} W_{N_0,N_1}(i,j) Q_{0j}] \tag{6}$$

In nearest neighbors matching, $Q_{1i} = Y_{1i}, Q_{0j} = Y_{0j}$, defined a matched sample as $C(X_i) = \parallel X_i, X_j \parallel < \varepsilon$, $W_{N_0,N_1}(i,j) = 1$ for matched observations, others are zero.

In kernel matching, the weights are $W_{N_0,N_1}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}$.

A regression adjusted matching has been complemented into DID by Heckman et al (1997). In this setting, $Q_{1i} = Y_{1it} - Y_{oit'}, Q_{0j} = Y_{0j} - Y_{0jt'}$

$$W_{N_0,N_1}(i,j) = \frac{G_{ij}}{\sum_{k \in I_0} G_{ik}}$$

A conditional DID matching estimator is been proposed as well,

$Q_{1i} = [(Y_{1it} - X_{it}\beta_{0t}) - (Y_{oit'} - X_{it'}\beta_{0t})]$=[($Y_{1it} - Y_{oit'}$)-($X_{it'} - X_{it}$) β_{0t}]

$Q_{1j} = [\left(Y_{1jt} - X_{jt}\beta_{0t}\right) - \left(Y_{ojt'} - X_{jt'}\beta_{0t}\right)]$=[($Y_{1jt} - Y_{ojt'}$)-($X_{jt'} - X_{jt}$) β_{0t}]

The setting is perfectly suitable for solving the problem of changes in characteristics of treatment group before and after time t. Chen and Jin (2012) suggest that the heterogeneity within a group allow us to control for the unobservable attributes based on an assumption that individual's unobservable attributes have the same distribution with observable attributes.[85] Halla and Zweimueller (2013) suggest combining DID and matching may effectively eliminated biases caused by unobservable attributes in the presence of longitudinal data.

5.4 Results.

5.4.1 Heterogeneity among university graduates.

[85] Unlike Heckman et al (1997) and Heckman et al (1998) in which they use longitudinal data, here they use all households within each county with unequal probabilities to participate the program to account for the unobservable heterogeneity attributes.

Figure 5.6 shows the changes in the returns between university graduates and non-university graduates given the propensity scores of attending universities. From Figure 5.6, it is clear shows that individual's return is negatively correlated with propensity scores. Obviously the highly significant differences are due to the fact that the university graduates and the non-university graduates are rather different based on the propensity scores.[86]

<Figure 5.6 Here>

In order to capture the intuition of the heterogeneity, I show that the individual's characteristics on the basis of the propensity score of attending universities. Figure 5.7 shows that individuals with the lowest level of attending university have substantially lower numbers of A-levels, number of GCSEs and year of education. Returning students have a lower university attendance rate, along with more working experiences and fewer years of education.

<Figure 5.7 Here>

Table 5.1 shows a simple breakdown of the types of student. For the university graduates, the average education of the returning students' increases compared to the fresh students. Presumably, the new returning students would come from GCSE, A-level graduates, and HNC or equivalent qualifications. But the mean

[86] The lowest propensity scores may consist of individuals who have lower levels of education. The middle part of the figure consists of A-level graduates. The right part with the highest propensity scores may come from individual with postgraduates and higher levels of vocational education, such as NVQ, HND etc. That implies there should be many channels which need to be controlled.

numbers of A-level for the returning students who have a degree doesn't change. It suggests that those new returning university graduates may come from individuals whose have a higher full-time vocational education. Moreover, the average education doesn't change among the returning students who don't have a degree. The results suggest that there are more students who have more years of education went to universities as a result of the reform. Not only the average years of education, but also the number of universities graduate increase for both types of the students.[87]

<Table 5.1 Here>

<Figure 5.8 Here>

Table 5.2 shows the 2SLS results for the fresh students and the returning students separately on the basis of the periods of attending universities. The results suggest that the returning students may benefit around 20% after becoming university graduates compared to the fresh students. The results are similar to Devereux and Fan (2011) in which they run the sufficient robustness checks. My results are quite similar to theirs, only differentiated by returning students and fresh students. The results suggest that the returning students and the fresh students have different returns after obtaining a degree and the returning students tend to have higher returns than the fresh students among

[87] Figure 5.8 also shows that the proportional change for returning students who obtain a degree as a result of the reform. It shows the distribution of continuous years of education of returning students who went to universities.

males. The returning students may have clear purposes of attending the university and stronger personal abilities as result of years of working experience.

<Table 5.2 Here>

5.4.2 DID and MDID.

To testify the source of bias, I split the sample into two groups, fresh students and returning students. Those individuals could be very diverse. Unlike the fresh students who go to university when they finish A-levels or higher levels of vocational education, returning students may have a more diverse educational background and working experiences. The QLFS dataset is not very informative regarding this perspective. It only includes the number of A-levels without the grades of those A-levels which does not allow me to capture the compositional change as a result of the reform for fresh university graduates. But it can capture those individuals who go to university with a lower number of A-levels or even without A-levels. In the Understanding Society, parental occupations are used to capture the compositional change.

Table 5.3 presents the DID and MDID results for the fresh students and the returning students respectively on the basis of different education expansion periods. Propensity Score Matching (PSM) and Coarsened Exact Matching (CEM) are used to balance the compositional changes. The DID results don't show significantly negative effects on the returns to the university graduates in both periods. For the MDID results, the most robust negative effects concentrate on the fresh students in the post-expansion period. The results are consistent in general between QLFS and Understanding Society. The negative effects only appearing in the post-expansion period may suggest that the quantitative effect is increasing with the increasing numbers of university graduates. That fact implies that we can improve the average years of education among the population with very little cost if we can subtly control the increase in the numbers of university graduates. The results also suggest that oversupply graduates don't decrease the returns of returning students, showing that the returning students have more stable returns. That may be because the returning students tend to have clear purposes before deciding to pursue a degree. In order to add the covariates to control for, I perform the DID strategy with weights, generated by the DID-PSM strategy. The results are shown in the brackets. After including more control variables in PSM, all results attenuate. The results between PSM and SEM show a very similar pattern and magnitude. With more covariates included in the regression such as subjects, the heterogeneities among the graduates could be better controlled.

<Table 5.3 Here>

Figure 5.9 shows the balance in father occupation in Understanding Society. This

figure is used to testify this condition $E(Y_{0t} - Y_{ot'}|P(Z), D = 0) - E(Y_{0t} - Y_{ot'}|P(Z), D = 1)$. Visually, CEM does a better job compared to PSM. Each bar presents the difference in the proportion of the university graduates and the non-university graduates under the father's occupations before and after the reform. Blue bars represent proportional changes in the proportion of university graduates given the father's occupations. Orange bars represent the proportional changes after weighted by SEM. Gray bars represent the proportional changes after weighted by PSM. Figure 5.10 shows the balance in number of A-levels. Figure 5.11 shows the balance in education for returning students. Clearly, CEM does a better work in the sense of balancing.

<Figure 5.9-5.11 Here>

5.4.3 Sensitivity test.

In order to show the robustness of the results, I perform the sensitivity test on the basis of the relative new technique. Oster (2016) proposed a new sensitivity test in which she argues that R-squared should take into consideration since the coefficients would not change massively when an uninformative control is included, as well as the R-squared. She proposed a method to derive a range from a controlled treatment effect to an unbiased treatment effect and take the R-squared into consideration compared to the Altonji, Elder and Taber (2005).

Table 5.4 shows the sensitivity of the CEM-DID results using the QLFS. In Table 5.4, the matching variables are different in the columns. The variables are all pre-treatment variables which are the variables of individuals before they went to university. I perform Oster's (2016) sensitivity test for the significant results in Table 5.3 in order to show the robustness of the results. The maximum of R-squareds are assumed to be 1, twice the R-squared_tilda and 1.25 times the R-squared_tilda. The R-squared_tilda is the R-squared of a fully controlled regression. In each weighted regression, the control variables are the same. Only the weighting variables are different. I check the sensitivity of the results only in post-expansion period since the results are seemingly more significant in Table 5.3. The QLFS-CEM-DID results show the strong robustness of the results. The ranges of the true treatment effect almost all lay within two standard deviations no matter how large the R-squared was.

Table 5.5 shows the sensitivity test of the CEM-DID results in Understanding Society. The results don't have any significant results on the basis of multiple sets of matching variables and R-squared. Although I notice that there is a significant positive effect in the third column, but the range of the treatment effect vary massively even when the maximum R-squared is very close to the R-squared_tilda.

<Table 5.4 and 5.5>

5.5 Conclusions.

Although the effect of the higher education expansion on returns has been largely examined, the results are still ambiguous to some extent due to measurement error and compositional changes. In this study, I apply MDID to examine the heterogeneous returns to the university graduates and highlight the differences in the return between the fresh students and the returning students.

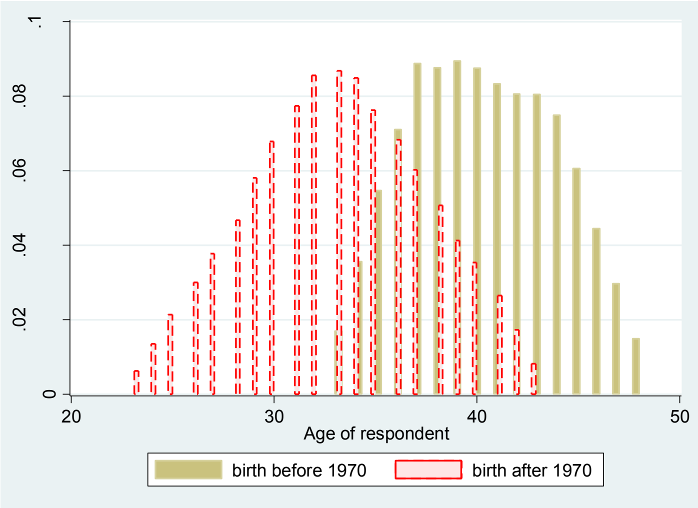
There are amount of returning students who obtain a degree as a result of the education expansion and the returns differ from the fresh students. The DID results are consistent with previous studies in that there are no significantly negative effects to the newly recruited university graduates. However, the MDID results correct the compositional change due to the fact that new graduates may have different educational backgrounds. The results suggest that there exists a significantly negative effect for the fresh students for both PSM-DID and SEM-DID results. The negative effects are more robust in the post-expansion periods. The negative effects concentrate on the fresh students. This might be due to the fact that there are more students with worse backgrounds becoming university graduates in the post-expansion period. The sensitivity tests show the negative effects for those fresh students are rather robust even the R-squared increases to one.

Given the results, there are enough reasons to believe that it may be not the best idea to push more students into universities. Individuals who have “worse” educational backgrounds and are not prepared to become university graduates need time and experiences to think the reasons before going into a university. Additionally, university level of education might be a potentially efficient path for some low-achievers to obtain more education and enhance their skills to make them more productive given the fact that they don’t have the university education due to multiple reasons.

This chapter has several limitations as a result of the data. Although the returning students may benefit more after attending the universities, the definition of the returning students is ambiguous due to its nature, such as the reasons of being a returning student, years of working experiences, measurement errors and so on. A data with richer information regarding the employment history could be very helpful to understand the mechanisms by looking inside subgroups of the returning students. Moreover, there is lack of information regarding the academic backgrounds of A-level graduates. The information regarding the background may bring great benefit to improve the matching results.

Figures and Tables:

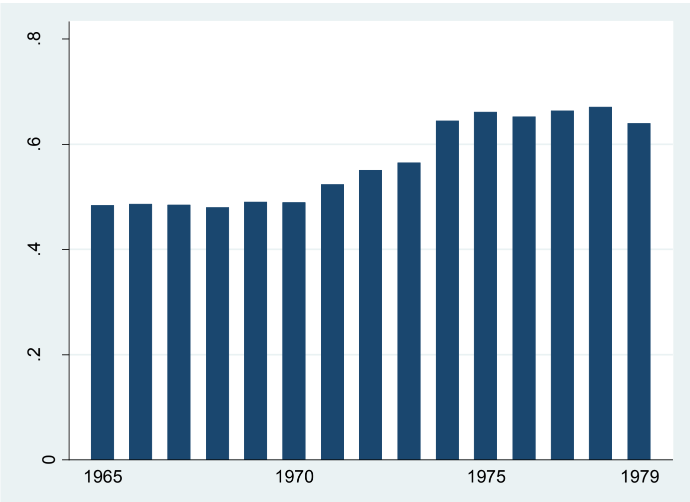
Figure 5.1 Proportion of age band before and after the reform



Notes: Proportion of observations born in different years, divided by born before and after 1970.

Sources: QLFS

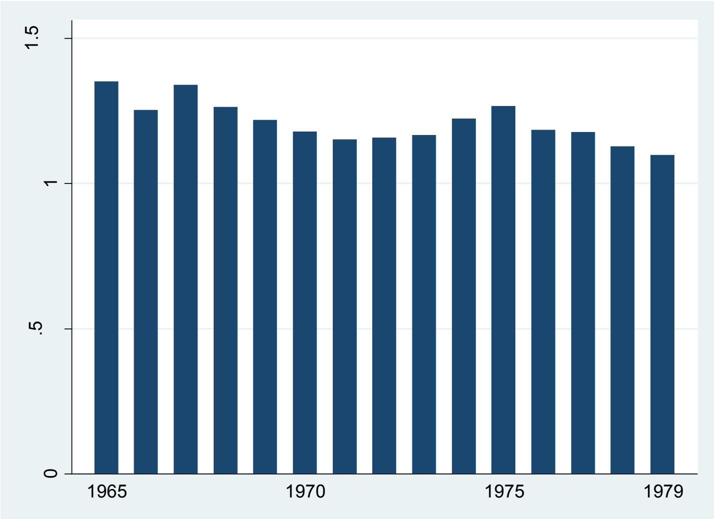
Figure 5.2 Number of A-levels given birth cohorts



Notes: Mean of average numbers of A-level.

Sources: QLFS

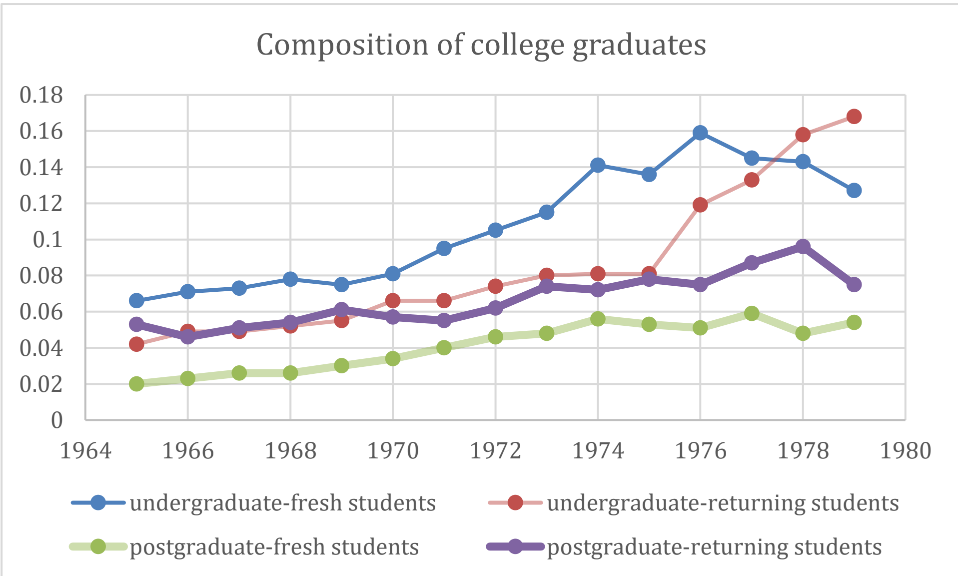
Figure 5.3 Number of A-levels among graduates given birth cohorts



Notes: Mean of average numbers of A-level among graduates.

Sources: QLFS

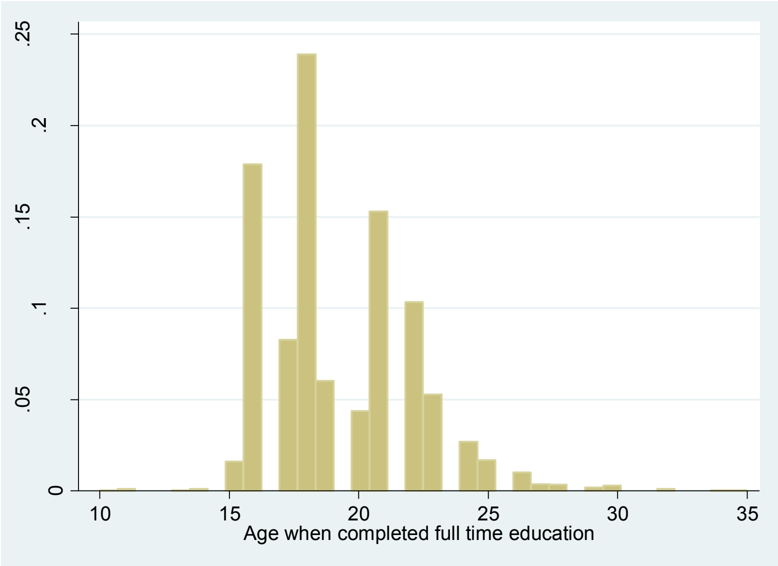
Figure 5.4 Composition of degree given birth cohorts



Notes: Y-axis represents the proportion of graduates among all qualifications on the basis of birth cohorts. The sample includes all observations.

Sources: QLFS

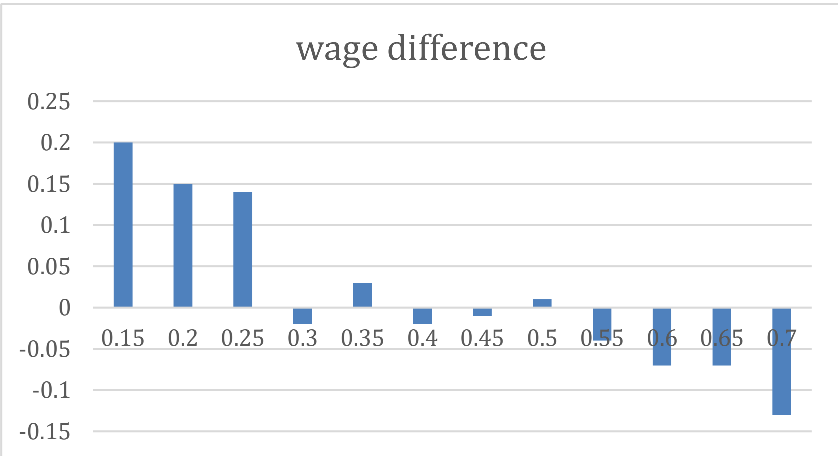
Figure 5.5 Ages when completed full-time education of returning students among university graduates.



Notes: Fraction of age when leaving continuous education among returning university graduates.

Sources: QLFS

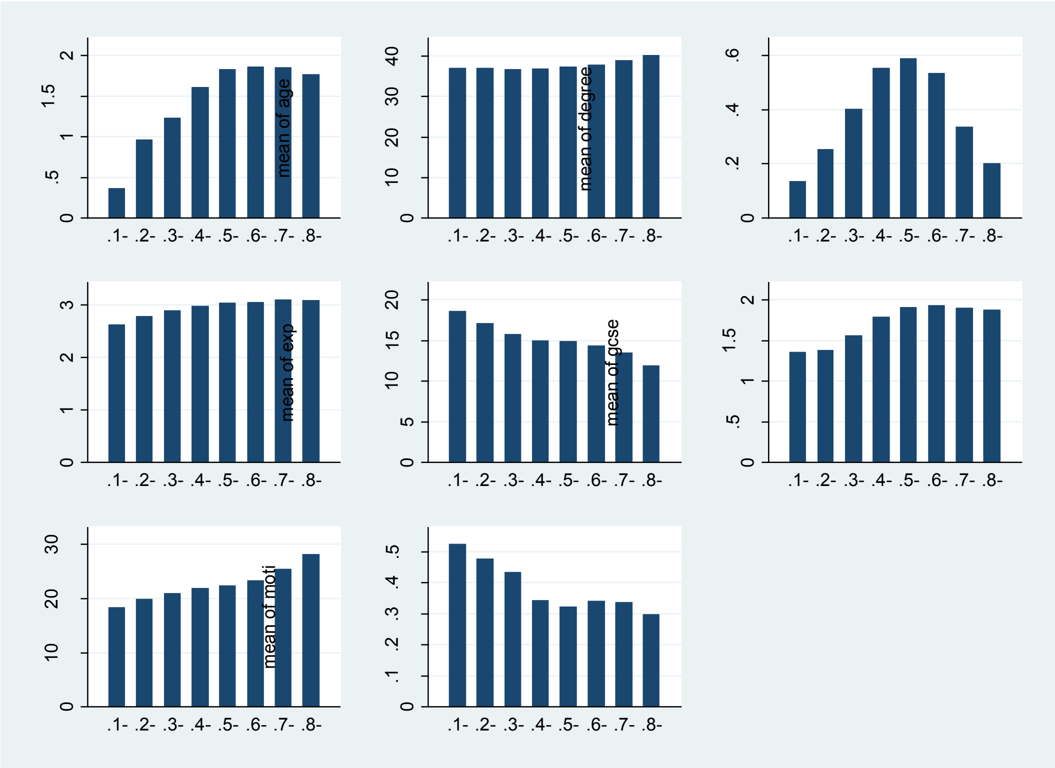
Figure 5.6 Wage difference between university graduates and non-university graduates



Notes: Sample periods includes birth cohort from 1965 to 1979. The wage differences are calculated by subtracting the real log of wage between university graduates and non-university graduates on the basis of propensity scores.

Sources: QLFS

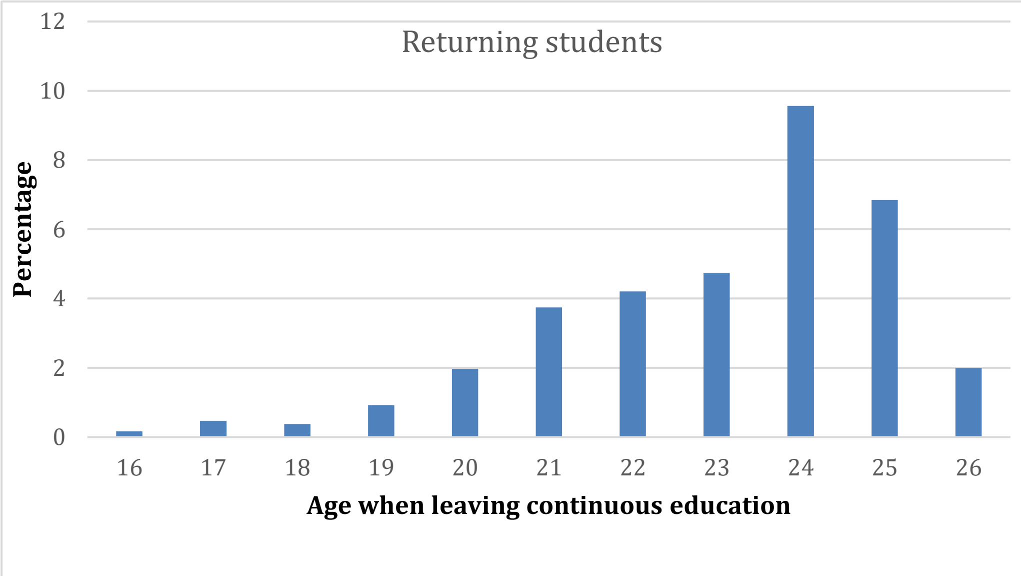
Figure 5.7 Characteristics given propensity score of degree



Notes: X-axis is propensity score of attending university. Here the propensity scores have been cut into bands. If squared of years of education is included, then the rate of degree would be linear given the propensity score.

Sources: QLFS

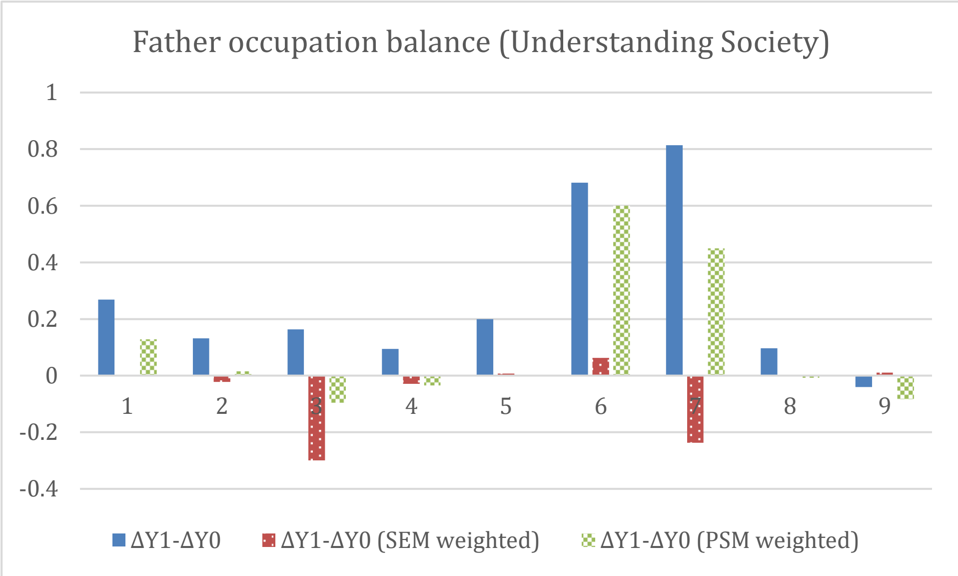
Figure 5.8 Proportional change for returning students obtaining degree before and after the reform



Notes: Relative proportional changes before and after the education expansion for returning students on the basis of years of full-time education. Y-axis is number of the proportional change compared to before the reform.

Sources: QLFS

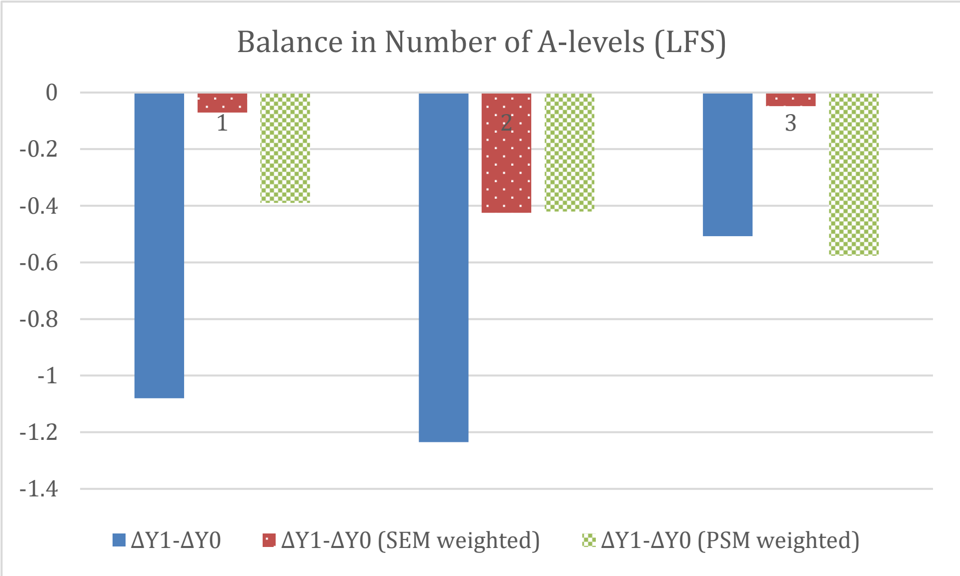
Figure 5.9 Balance of father occupation



Notes: ΔY1 represents the differences in the numbers of the university graduates before and after the reform given father's occupation. ΔY0 represents the non-university graduates. The blue bars identify the relatively proportional change between university graduates and non-university graduates. The orange bars represent the situation after re-weighted by SEM. The grey bars are after PSM weighted. Visually the SEM works better than PSM.

Sources: Understanding Society

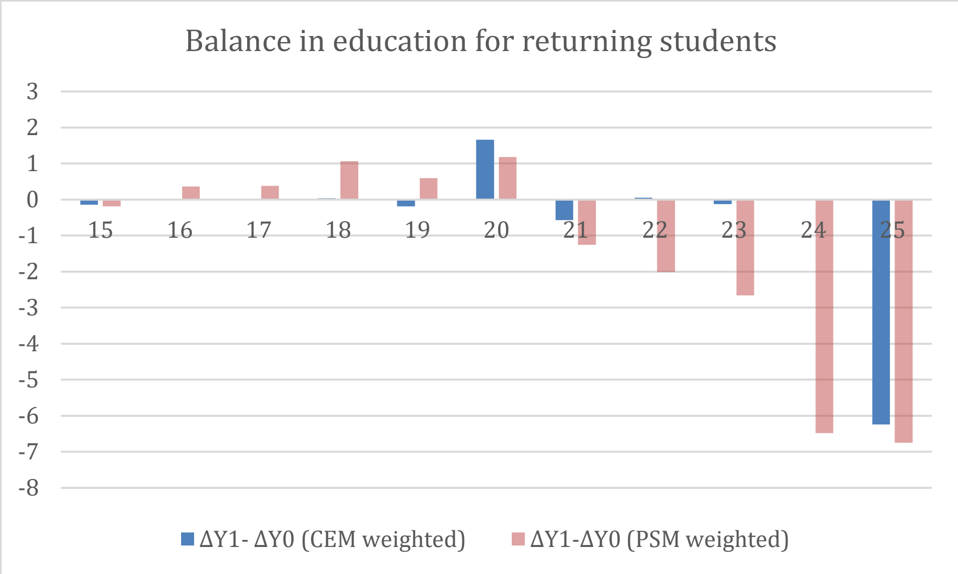
Figure 5.10 Balance in Number of A-levels (QLFS)



Notes: This figure is similar to the previous figure. It describes the balance in number of A-levels. "1", "2", and "3" represent none, one, and two and above two A-levels respectively.

Sources: QLFS

Figure 5.11 Balance in education for returning students



Notes: Above figure describes the balance in the years of education for returning students since there are rather different results between SEM weighted and PSM weighted. Visually, the PSM has failure in capturing the relatively proportional change in the years of education caused by the education reform compared to the SEM.

Sources: QLFS

Table 5.1 Statistic summary

Undergraduate	Before reform (Born<1970)		After reform (Born>=1970)	
	Fresh students	Returning students	Fresh students	Returning students
Sample size	2386	1626	4428	3306
Education	22.02	18.49	22.14	20.16
Experience	20.36	21.53	16.40	18.72
No. A-level	1.54	0.898	1.46	0.895
No. GCSE	0.898	0.844	0.881	0.728
Non-Undergraduate				
Sample size	15974	12620	15126	15875
Education	17.68	17.34	18.55	17.70
Experience	20.94	21.67	17.19	19.13
No. A-level	0.370	0.376	0.476	0.386
No. GCSE	0.715	0.619	0.731	0.638

Notes: The above panel summarizes individual's characteristics among university graduates and below panel summarizes individual's characteristics among non-university graduates. "No. A-level" and "No. GCSE" are categorical variables which indicate rough numbers of A-level or GCSE.

Sources: QLFS

Table 5.2 2SLS results

		Fresh students	Returning students
2SLS			
1965-1975			
➔	Male	0.051***	0.060***
➔	Female	0.063***	0.059***
1965-1969 and 1976-1979			
➔	Male	0.056***	0.062***
➔	Female	0.066***	0.061***

Notes: The 2SLS results are similar results with Devereux and Fan (2011). I only separately run the regression based on sample periods and types of students.

Sources: QLFS

Table 5.3 DID and MDID between individuals with or without working experiences for undergraduates

	QLFS		Understanding Society	
	Fresh students	Returning students	Fresh students	Returning students
DID				
Pre-expansion	-0.013	-0.020	-0.081	0.000
Post-expansion	-0.056*	-0.008		
MDID (PSM)				
Pre-expansion	-0.058***	0.016	-0.126***	-0.010
	(-0.030*)	(0.009)	(-0.086**)	(0.014)
Post-expansion	-0.129***	-0.000		
	(0.083***)	(0.001)		
MDID (CEM)				
Pre-expansion	-0.028	-0.011	-0.098	0.003
Post-expansion	-0.065***	-0.038		

Notes: Given the result of female without working experiences, there might be other factors for which those could bias the results. The results only include males. In QLFS, the control variables are numbers of A-level, numbers of GCSE, experiences, squared of experiences, years of tenure, have job training, disability, marriage, work in London, full-time job, quarter, year, industry and subject in university. In Understanding Society, the control variables are age, years of education, marriage, father occupation, year and industry.

Sources: QLFS and Understanding Society

Table 5.4 Sensitivity test for QLFS

Estimates	(1)	(2)	(3)
Birth cohorts	+	+	+
A-levels		+	+
GCSEs			+
CEM			
Post-expansion (Fresh)	-0.073**	-0.065**	-0.065***
R squared	0.344	0.306	0.289
N	9361	10008	9249
PSACALC (Rmax=1)	[-0.115, -0.073]	[-0.092, -0.063]	[-0.065, 0.070]
PSACALC (Rmax=2*R squared)	[-0.091, -0.073]	[-0.075, -0.063]	[-0.065, -0.018]
PSACALC (Rmax=1.25*R squared)	[-0.077, -0.073]	[-0.067, -0.063]	[-0.065, -0.054]
Post-expansion (Mature)	-0.031	-0.038	-0.021
R squared	0.336	0.344	0.28
N	8677	8030	7998
PSACALC (Rmax=1)	[-0.031, 0.035]	[-0.037, 0.129]	[-0.022, 0.157]
PSACALC (Rmax=2*R squared)	[-0.031, 0.139]	[-0.037, 0.039]	[-0.022, 0.043]
PSACALC (Rmax=1.25*R squared)	[-0.031, 0.007]	[-0.037, -0.020]	[-0.022, -0.006]

Notes: Solely for the male. There is a trade-off between bias and efficiency for CEM since the strata could be enormous with increasing variables. PSM may not encounter this problem. For CEM, the weights are estimated with different covariates. Birth cohorts, A-levels, GCSEs, education, marriage, and disable are pre-treatment variables. “Rmax” in the PSACALC test is 1 which is the R squared.

Sources: QLFS

Table 5.5 Sensitivity test for Understanding Society

Estimates	(1)	(2)	(3)
Birth cohorts	+	+	+
Father occupation		+	+
Mother occupation			+
CEM			
Pre-expansion (Fresh)	-0.087	-0.098	-0.061
R squared	0.306	0.309	0.307
N	4729	4397	3222
PSACALC (Rmax=1)	[-0.087, 5.76]	[-0.097, 5.68]	[-26.6, -0.061]
PSACALC (Rmax=2*R squared)	[-0.087, 3.79]	[-0.097, 3.76]	[-26.0, -0.061]
PSACALC (Rmax=1.25*R squared)	[-0.087, 1.85]	[-0.097, 1.83]	[-25.6, -0.061]
Pre-expansion (Mature)	0.002	0.003	0.078**
R squared	0.250	0.226	0.230
N	7344	7587	6305
PSACALC (Rmax=1)	[-2.78, 0.002]	[0.003, 2.84]	[0.078, 1.88]
PSACALC (Rmax=2*R squared)	[-1.60, 0.002]	[0.003, 1.53]	[0.078, 0.494]
PSACALC (Rmax=1.25*R squared)	[-0.80, 0.002]	[0.003, 0.766]	[0.078, 0.171]

Notes: All results are based on the observations after dropping the missing observations of father's occupation. The sample decreases around 30% after dropping the missing mother's occupation observations. Including the missing observations doesn't make significant change to the results only making the variance larger. I suspect that there is serious multicollinearity among parental information.

Sources: Understanding Society

Chapter 6

General Conclusions

6.1 Findings.

This thesis aims to address the economic outcomes of education regarding some crucial individuals' education decisions. In the first empirical chapter, I focus on the returns to vocational education. The return to vocational education exists over time compared to academic education over time. The returns also vary by the types of vocational education. Although the returns to vocation education might be biased due to the age effect, it describes the relatively differences in returns between types of education over birth cohorts. Vocational education tends to have less promising career development in the future and the returns to vocational education become smaller in recent years. Moreover, the education expansion has negative effects on the returns to vocational education. The negative effects are smaller in the during-expansion periods when there are fewer graduates compared to post-expansion periods, suggesting that the quantitative effect of increasing graduates reduce the returns to vocational education.

In general, the results suggest that vocational education leads to lower earning compared to academic education in the middle of career. The differences come from different educational background, personal abilities. The DID results suggest that the returns to vocational education are still affected by the over-supply of academic students with similar levels of education. This chapter complements the literature in terms of the effects of education expansion on vocational education. It implies that the overall effect of increasing university graduates on the society might be underestimated if it only counts the negative effects on university graduates.

In the second chapter, I examine the effect of becoming eligible for adult minimum wage rate on employment probability for different groups. There is a strong evidence of heterogeneous effects by qualifications. The results suggest that individuals with higher numbers or grades of GCSE have a higher probability of being employed after increasing the minimum wage. Besides the general employment opportunity, the results strongly suggest that higher skilled workers tend to find a more full-time permanent or full-time job after becoming eligible for higher minimum wage rate. But on the other hand, lower skilled workers have a lower probability of finding a full-time permanent job. The evidence of higher employment probability and higher satisfying job accession probability may imply that there is a crowding out effect coming from higher skilled workers. Although the result is modestly significant, the higher number of GCSEs and the higher proportion of 5+ GCSE suggest the existence of crowding out effect directly. On the other hand, there is no evidence in terms of lower skilled workers being made redundant. However, the limited information of the dataset does not allow me to examine the replacement effect in more details or labour flows in general. With data in a booming period, one might expect more variation between the local unemployment rate and the employment probability.

The implication of this chapter is that we should not neglect the possible adverse effects of the increase in minimum wage, especially in a recessionary period. Due to a tighter labour market, the discontinuity caused by increasing minimum wage will not only increase labour supply but also result in less chance for disadvantaged workers in labour market. Moreover, even though the negative effect of minimum wage is limited on average, the negative effect in subgroups may be still non-negligible. The results imply that exogenous increase in competition may have both immediate and long-term negative effects on disadvantaged workers. The minimum wage policy should be more flexible in a tighter labour market. The RD design could be more flexible and convincing if day of birth is available to construct the distances. Moreover, the stock of employment could be ambiguous. It might be clearer to examine crowding out effect on the basis of examining flow of labour, but that requires a more comprehensive dataset.

In the third chapter, I apply MDID to examine the heterogeneous returns to the university graduates and highlight the differences in returns between the fresh students and the returning students. The DID results are consistent with the previous studies that there are no significantly negative effects to the newly recruited university graduates. However, the MDID results correct the bias due to the fact that new graduates may have lower innate ability. It shows significantly negative effects for the fresh students. The results are consistent on the basis of the different matching strategies. Both matching strategies examine a strong negative effect for the fresh students in the post-expansion period. That might be due to the fact that there are more students with worse backgrounds becoming university graduates in the post-expansion period.

Given the results, there are enough reasons to believe that it may be not the best idea to push more fresh graduates into universities. Getting them into employment is hard and pushing them into education is easy, but individuals who have "worse" background and are not prepared to become university graduates need time and experiences to think the reason for going into a university.

6.2 Further work.

The returns to education has been both extensively and intensively examined.[88] The selection between vocational and academic education has been a stubborn

[88] Vocational education is more versatile than academic education. It serves multiple purposes in the society. The purposes regarding vocational education are also various in different countries. Due to its nature vocational education tends to be lack of general training in terms of cognitive skills and non-cognitive skills. Countries like U.S. do not emphasize the importance of vocational education in the level of secondary education. On the other hand, countries like Germany have long tradition regarding vocational education in the labour market.

problem in the literature. Most studies prefer to rely on a rich dataset which includes a comprehensive description regarding the early academic achievements and family backgrounds. However, due to the complex self-selection and the nature of multi-dimensional personal abilities the results might be still ambiguous. A more promising method is to take advantage of the quasi-experiment which converts vocational education into academic education or verse visa. Moreover, the disparities between academic and vocational tracks for the students might be existing since the beginning. The effects of early educational attainment on the selection of vocational education may be worth exploring. Intergenerational mobility has been both intensively and extensively examined. The evidence has shown that parents have influence on children's education attainment, future income, mental health and so on. A consensus about interventions to children's education is that early intervention is both effective and efficient. However, it is unrealistic to change the parental status. Primary education is children's first formal and regulated education in their life. Practically, primary education can be considered as a promising intervene for the government to implement. Besides, little is known regarding the relation between early educational achievement and future educational achievement. The major difficulty is that there might be too many channels to pin down. The results would be ambiguous if only depending on the control methods.

In the two empirical chapters, the return to education normally leaves a much unexplained part in the standard Mincer's equation, even on the basis of rich datasets regarding an individual's background.[89] And in the recent decades, there is a growing income inequality in the U.S. and UK. The analysis of income inequality across age, education, and occupation have been carried out intensively. The contribution of within-group inequality to the overall income inequality has been neglected (Barth et al, 2016). This motivates me that the firm's characteristics may explain the within-education cell inequality. There are comprehensive evidence argues that firm-size has significant impacts on the wages differentials (Troske, 1999; Balkan and Tumen, 2016, etc). Lindley and McIntosh (2015) examine the income inequality among university graduates given variance analysis. The similar extension of the variance analysis could be performed on the basis of firm-level or industry-level, as well as the variance decomposition proposed by Lemieux (2006). Labour Force Survey (LFS) is a comprehensive dataset which describes the labour market and it covers most of the countries in the Europe. British Cohort Study (BCS) and National Child Development Study (NCDS) might be also helpful as it includes rich information regarding the early educational attainments. However, an employer-employee data would be helpful to pin down the channels of heterogeneous firm impacts in this study, such as Income Data Services (IDS) in the UK. The contribution of this

[89] Although the employment process involves randomness, some of the potential factors that we may neglect are the factors of the workplaces. Two identical workers may have different returns in because of the firm's characteristics, such as physical capital intensity (skill and capital complementary), market power (higher profitability), monitoring, unionization, workers clusters (highly productive workers in a team).

study would be twofold. First, by discovering the firm characteristics on the income inequality, we may have a better understanding why the standard Mincer wage equation leaves the unexplained error term. Second, with rich industry-level data we may acknowledge how firm characteristic affects income inequality.

There is a widely accepted argument that minimum wage has influential impacts on the level of employment. However, the change in the level of employment may have multiple reasons, such as job accession rate, job separation rate, turnover rate, and so on. Recent literature regarding minimum wage focuses on the flow of labour. There is few study examining the flow of labour in the Europe, due to the limit of the data. The existing literature regarding the flow of labour is based on the establishment-level dataset. Moreover, the minimum wage of apprenticeship was introduced in 2010. It may encourage young workers to convert the career path. After realizing the effects of increasing the minimum wage on the level of employment and the flow of labour for the younger workers, it is interesting to know how to escape from the lower return jobs. Many literature has suggested that there exists strong low-pay-no-pay cycle among lower level jobs. After getting older and having more working experiences, the younger workers may want to find better paid jobs. Although the literature suggests the state dependence, the mechanisms have not been discovered. Do the signals sending from a low-paid job make them less likely end in a better paid job? Or the previous jobs simply don’t accumulate the useful human capital for the future career? Finding the reason behind the state dependence becomes very important for the individuals and policy makers to regulate the education policies.

References:

Abadie, Alberto. 2005. Semiparametric Difference-in-Differences Estimators. The *Review of Economic Studies* 72 (1): 1-19.

Addison, John T., McKinley L. Blackburn, and Chad D. Cotti. (2013) Minimum wage increases in a recessionary environment. *Labour Economics* 23: 30-39.

Altonji, Joseph G, Todd E. Elder., and Christopher R. Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113(1): 151-184.

Andersson, Roland., Pardis Navavi, and Mats Wilhelmsson. 2014. The impact of advanced vocational education and training on earnings in Sweden. *International Journal of Taining and Development* 18(4): 256-270.

Angrist, Joshua., and Jinyong, Hahn. 2004. When to control for covariates? Panel asymptotics for estimators of treatment effects. *Review of Economics and Statistics* 86(1): 58-72.

Angrist, Joshua D., and Joern-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Angrist, Joshua D., and Alan Krueger. 1991. Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics*, 106(4):

979-1014.

Angrist, Joshua D., and Miikka Rokkanen. 2015. Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff. *Journal of American Statistical Association* 110(512): 1331-1344.

Arulampalam, Wiji., Alison L. Booth, and Mark, L. Bryan. (2004) Training and the New Minimum Wage. *The Economic Journal* 114(494): c87-c94.

Ashenfelter, Orley., and Card, David. 1985. Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics* 67 (4): 648-660.

Balkan, Binnur., and Semih Tumen. 2016. Firm-Size wage gaps along the Formal-Informal divide: theory and evidence. *Industrial Relations* 55(2).

Backes-Gellner, Uschi., and Regula Geel. 2014. A comparison of career success between graduates of vocational and academic tertiary education. *Oxford Review of Education* 40(2): 266-291.

Barth, Erling., Alex Bryson, James C. Davis., and Richard Freeman. 2016. It's where you work: Increases in the dispersion of earnings across establishments and Individuals in the United States. *Journal of Labour Economics* 34(2).

Bertrand, Marianne, Esther Duflo., and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of*

Economics 119 (1): 249-275.

Billett, Stephen. 2011. *Vocational Education: Purposes, Traditions and Prospects*. Dordrecht Heidelberg London New York: Springer.

Birch, Elisa Rose., and Paul W. Miller. 2007. The Characteristics of 'Gap-Year' Students and Their Tertiary Academic Outcomes. *The Economic Record* 83(262): 329-344.

Bishop, John H., and Ferran Mane. 2004. The impacts of career-technical education on high labour market success. *Economics of Education review* 23(4): 381-402.

Bhuller, Manudeep., Magne Mogstad, and Kjell G. Salvanes. 2017. Forthcoming. Life Cycle Earnings, Education Premiums and Internal Rates of Return. *Journal of Labour Economics.*

Blundell, Richard., Lorraine Dearden, Alissa Goodman, and Howard Reed. 2000. The Returns to Higher Education in Britain: Evidence from a British Cohort. *The Economic Journal* 110(461): F82-F99.

Blundell, Richard., Lorraine Dearden, and Barbara Sianesi. 2005. Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of Royal Statistics Society A* 168(3): 473-512.

Blundell, Richard., and Monica Costa Dias. 2009. Alternative Approaches to Evaluation in Empirical Microeconomics. *Journal of Human Resources* 44(3):

565-640.

Booth, Alison L., Marco Francesconi, and Jeff Frank. 2002. Temporary Jobs: Stepping Stones or Dead Ends? *The Economic Journal* 112(480): 189-213.

Brochu, Pierre., and David A. Green. 2013. The impact of minimum wages on labour market transitions. *The Economic Journal* 123(573): 1203-1235.

Card, David., and Alan B. Krueger. 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review* 84(4): 772-793.

Carneiro, Pedro., James J. Heckman, and Edward J. Vytlacil. 2011. Estimating Marginal Returns to Education. *American Economic Review* 101(6): 2754-2781.

Castex, Gonzalo., and Evgenia Kogan Dechter. 2014. The Changing Roles of Education and Ability in Wage Determination. *Journal of Labour Economics*, 32(4): 685-710.

Chen, yuyu., and Jin Ginger Zhe. 2012. Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. *Journal of Health Economics* 31(1): 1-14.

Chowdry, Haroon., Claire Crawford, Lorraine Dearden, Alissa Goodman, and Anna Vignoles. 2010. Widening Participation in Higher Education: Analysis using Linked Administrative Data. *IFS Working Paper:* W10/04.

Conlon, Gavan. 2005. The Determinants of Undertaking Academic and Vocational

Qualifications in the United Kingdom. *Education Economics* 13(3): 299-313.

Cuddy, Natalia., and Tom Leney. 2005. Vocational education and training in the United Kingdom. Cedefop Panorama series; 111. European Centre for the Development of Vocational Training (Cedefop).

Dearden, Lorraine. 1999. The effects of families and ability on men's education and earnings in Britain. *Labour Economics* 6(4): 551-567.

Dearden, Lorraine., Steven Mcintosh, Michal Myck, and Anna Vignoles. 2002. The Returns to Acedemic and Vocational Qualifications in Britain. *Bulletin of Economic Research* 54(3): 0307-3378.

Dearden, Lorraine., Carl Emmerson, Christine Frayne, Alissa Goodman, Hidehiko Ichimura, Costas Meghir, Karl Ashworth, Jay Hardman, Woon-Chia Liu, Sue Maguire, and Sue Middleton. 2001. Education Maintenance Allowance: The First Year A Quantitative Evaluation. Department for Education and Employment, Research Report, No. 257.

Dearden, Lorraine., Howard Reed, and John Van Reenen. 2006. The impact of Training on Productivity and Wages: Evidence from British Panel Data. *Oxford Bulletin of Economics and Statistics* 68(4): 0305-9049.

Dearden, Lorraine., Emla Fitzsimons, and Gill Wyness. 2011. The Impacts of Tuition Fees and Support on University Participation in the UK.

Dearden, Lorraine., Alissa Goodman, and Gill Wyness. 2012. Higher Education Finance in the UK. *Fiscal Studies* 33(1): 73-105.

Devereux, Paul, J., and Robert, A. Hart. 2010. Forced to be rich? Returns to compulsory schooling in Britain. *The Economic Journal*, 120(549): 1345-1364.

Devereux, Paul J., and Wen Fan. 2011. Earnings returns to the British education expansion. *Economics of Education Review* 30(6): 1153-1166.

Dickens, Richard., and Alan Manning. 2004. Spikes and spill-overs: The impact of the national minimum wage on the wage distribution in a low-wage sector. *The Economic Journal*, 114(494): C95-C101.

Dickens, Richard., Rebecca Riley, and David Wilkinson. 2014. The UK minimum wage at 22 years of age: a regression discontinuity approach. *Journal of Royal Statistical Society* 177(1): 95-114.

Dickens, Richard., Rebecca Riley, and David Wilkinson. 2015. A Re-examination of the Impact of the UK National Minimum Wage on Employment. *Economica* 82(328), 841-864.

Dolado, Juan J., Florentino Felgueroso, and Juan F. Jimeno. 2000. Youth labour markets in Spain: Education, training, and crowding-out. *European Economic Review* 44(4): 943-956.

Dolton, Peter., and Anna Vignoles. 2002. Is a broader curriculum better?

Economics of Education Review 21(5): 415-429.

Dong, Yingying. 2016. Jump or Kink? Regression Probability Jump and kink Design for Treatment Effect Evaluation. University of California Irvine.

Draca, Mirko, Stephen Machin., and John Van Reenen. 2011. Mininum Wages and Firm Profitability. *American Economic Journal: Applied Economics* 3(1): 129-151.

Dube, Arindrajit., T. William Lester, and Michael Reich. 2016. Minimum Wage Shocks, Employment Flows, and Labour Markets Frictions. *Journal of Labour Economics* 34(3): 663-704.

Engellandt, Axel., and Regina T. Riphahn. 2005. Temporary contracts and employee effort. *Labour Economics* 12(3): 281-299.

Falk, Armin., Renst Fehr, and Christian Zehnder. 2006 Fairness perceptions and researvation wages – The behavioral effects of minimum wage laws. *Quarterly Journal of Economics* 121(4): 1347-1381.

Fidrmuc, Jan., and J.D. Tena. 2009. National minimum wage and employment of young workers in the UK. Department of Economics and Finance, Brunel University London.

Fuller, Alison. 2016. The Growth of Apprenticeship in England: Doubts Beneath the Numbers. *Challenge* 59 (5): 422-433

Fuller Alison., and Lorna Unwin. 2003. Creating a 'Modern Apprenticeship': A critique of the UK's multi-sector, social inclusion approach. *Journal of Education and Work* 16(1): 5-25.

Further Education Funding Council. 1994. *National Vocational Qualifications in the Further Education Sector in England*. Coventry.

Freeman, Richard D. 1976. The overeducated American. New York: American press, c1976.

Gautier, Pieter A., Gerard J. van den Berg, Jan C. van Ours, and Geert Ridder. 2002. Worker turnover at the firm level and crowding out of lower educated workers. *European Economic Review* 46(3): 523-538.

Glynn, Adam N., and Kevin M. Quinn. 2010. Introduction to the Augmented Inverse Propensity Weighted Estimator. *Political Analysis* 18(1):36-56.

Giuliano, Laura. 2013. Minimum Wage Effects on Employment, Substitution, and the Teenage Labour Supply: Evidence from Personnel Data. *Journal of Labour Economics* 31(1): 155-194.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. Identification and Estimation of Treatment Effects with a Regression-Discontinuity. *Econometrica* 69(1): 201-209.

Halla, Martin., and Martina Zweimueller. 2013. The effect of health on earnings: Quasi-expenrimental evidence from commuting accidents. *Labour Economics* 24:

23-38.

Hawley, Joshua D. 2003. Comparing the payoff to vocational and academic credentials in Thailand over time. *International Journal of Educational Development* 23(6): 607-625.

Hirano, Keisuke., and Guido W. Imbens. 2001. Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Searvices & Outcomes Research Methodology* 2(3-4): 259-278.

Heckman, James, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47(1): 153-161.

Heckman, James J., and Richard Robb. 1985. Alternative methods for evaluating the impact of interventions. *Journal of Econometrics* 66(1): 239-267.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies* 64(4): 605-654.

Heckman, James., Hidehiko Ichimura, Jeffrey Smith., and Petra Todd. 1998. Characterizing Selection Bias Using Experimental Data. *Econometrica* 66(5): 1017-1098.

Heckman, James J., and Edward Vytlacil. 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73(3): 669-738.

Heckman, James J, Hidehiko Ichimura., and Petra E. Todd. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies* 64(4): 605-654.

Hotchkiss, Lawrence. 1992. Effects of Training, Occupation, and Training-Occupation Match on Wage. The *Journal of Human Resources* 28 (3): 482-496.

House of Commons. 2016. Tuition Fee Statistics.

Hupkau, Claudia., Sandra McNally, Jenifer Ruiz-Valenzuela, and Guglielmo Ventura. 2016. Post-Compulsory Education in England: Choices and Implications. Centre for Vocational Education Research, Discussion Paper 001, ISSN: 2398-7553.

Huber, Martin, Michael Lechner., and Conny Wunsch. 2013. The performance of Estimators based on the propensity score. *Journal of Econometrics* 175(1): 1-21.

Hupkau, Claudia., and Guglielmo Ventura. 2017. Further Education in England: Learners and Institutions. Centre for Vocational Education Research, Briefing Note 001.

Iacus, Stefano M., Gary King., and Guiseppe Porro. 2012. Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Anlysis* 12 (1): 1-24.

Imbens, Guido W. 2015. Matching Mehods in Practice: Three Examples. The

Journal of Human Resources 50(2): 373-419.

Imbens, Guido W., and Joshua D. Angrist. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica* 62(2): 467-475.

Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practive. *Journal of Econometrics* 142(2): 615-635.

Kabatek, Jan. 2015. Happy Birthday, You're Fired! The Effects of Age-Dependent Minimum Wage on Youth Employment Flows in the Netherlands. *IZA* No. 9528.

Kahyarara, Godius., and Francis Teal. 2007. The Returns to Vocational Training and Academic Education: Evidence from Tanzania. *World Development* 36(11): 2223-2242.

Lechner Micheal., and Anthony Strittmatter. 2017. Forthcoming. Practical Procedures to Deal with Common Support Problems in Matching Estimation. *Econometric Reviews*.

Lee, David S. 2008. Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics* 142: 675-697

Lee, David S., and David Card. 2008. Regression discontinuity inference with specification error. *Journal of Econometrics* 142(2): 655-674.

Lee, David S., and Thomas Lemieux. 2009. Regression discontinuity designs in economics. National Bureau of Economic Research 14723.

Lemieux, Thomas. 2006. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skills? *The American Economic Review* 96(3): 461-498.

Lene, Alexandre. 2011. Occupational downgrading and bumping down: The combined effects of education and experience. *Labour Economics* 18(2): 257-269.

Lindley, Joanne., and Steven Mcintosh. 2015. Growth in within graduate wage inequality: The role of subjects, cognitive skill dispersion and occupational concentration. *Labour Economics* 37: 101-111.

Machin, Stephen., Lupin Rahman, and Alan Manning. 2003. Where the minimum wage bites hard: Introduction of minimum wages to a low wage sector. *Journal of the European Economic Association* 1(1): 154-180.

Machine, Stephen., and Anna Vignoles. 2005. *What's the good of education? : the economics of education in the UK*. Princeton University Press.

Malamud, Ofer., and Cristian Pop-Eleches. 2010. General evidence from an economy in transition. The *Review of Statistics and Economics* 92 (1): 43-60.

Martin, Daniel and Olivier Pierrard. 2014. On-the-job search and cyclical unemployment: crowd out vs. vacancy effects. *Journal of Economic Dynamic & Control* 44: 235-250.

Mcintosh, Steven. 2006. Further Analysis of the Returns to Academic and Vocational Qualifications. *Oxford Bulletin of Economics and Statistic*s, 68(2):

0305-9049.

Meer, Jonathan. 2007. Evidence on the returns to secondary vocational education. *Economics of Education Review* 26(5): 559-573.

Meghir, Costas., and Marten Palme. 2005. Educational Reform, Abiligy, and Family Background. *American Economic Review* 95(1): 414-424.

Metcalf, David. 2008. Why has the British National Minimum Wage had Little or No Impaction Employment? *Journal of Industrial Relations* 50 (3): 489-512.

Moenjak, Thammarak., and Christopher Worswick. 2003. Vocational education in Thailand: a study of choice and returns. *Economics of Education Review* 22(1): 99-107.

Neuman, Shoshana., and Adrian Ziderman. 1999. Vocational Education in Israel. *The Journal of Human Resources* 34(2): 407-420.

Nunez, Imanol., and Llias Livanos. 2015. Temps "by choice"? An Investigation of the Reasons Behind Temporary Employment Among Young Workers in Europe. *Journal of Labour Research* 36(1): 44-66.

Oster, Emily. 2017. Forthcoming. Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business and Economic Statistics*.

Polidano, Cain., and Domenico Tabasso. 2014. Making it real: The benefits of workplace learning in upper-secondary vocational education and training courses. *Economics of Education Review* 42: 130-146.

Prada, Maria F., and Sergio Urzua. 2017. Forthcoming. One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance and Earnings. *Journal of Labour Economics*.

Riley, Rebecca., and Chiara Rosazza Bondibene. 2017. Raising the standard: Minimum wages and firm productivity. *Labour Economics* 44: 27-50.

Rosenbaum, Paul, R., and Donald B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.

Stewart, Mark B. 2004. The impact of the introduction of the U.K. minimum w on the employment probabilities of low-wage workers. *Journal of the European Economic Association* 2 (1): 67-97.

Troske, Kenneth R. 1999. Evidence on the employer size-wage premium from worker-establishment matched data. *The Review of Economics and Statistics* 81(1): 15-26.

Trosel, Philip, Ian Walker, and Paul Woolley. 2002. Estimates of the economics return to schooling for 28 countries. *Labour Economics* 9(1): 1-16.

Van den Berg, Gerard J., and Geert Ridder. 1998. An Empirical Equilibrium Search Model of the Labour Market. *Econometrica* 66(5): 1183-1221.

Vignoles, Anna, Augustin De Coulon., and Oscar Marcenaro-Gutierrez. 2010. The value of basic skills in the British labour market. *Oxford Economic Papers* 63(1): 27-48.

Walker, Ian., and Yu Zhu. 2008. The College Wage Premium and the Expansion of Higher Education in the UK. *Scandinavian Journal of Economics* 110 (4): 695-709.

Walker, Ian., and Yu Zhu. 2011. Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review* 30(6): 1177-1186.

West., John, and Hilary Steedman. 2003. Finding Our Way: Vocational Education in England. *Centre for Economic Performance, London School of Economics and Political Science.*

Wooldridge, Jeffrey M. 2014. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50(2): 420-445.